

Differences in Word Usage by Truth Tellers and Liars in Written Statements and an Investigative Interview After a Mock Crime

DAVID MATSUMOTO* and HYISUNG C. HWANG

San Francisco State University and Humintell, LLC

Abstract

Although much is known about word usage differences between truths and lies, most of the research to date involves an examination of truths and lies in low stakes situations, written statements or interviews but not both, and native speakers of a single language. We examined differences in word usage between truth tellers and liars in a moderately high stakes, real-life scenario (mock crime) involving participants from four cultural/ethnic groups—European-Americans, Chinese, Hispanics and Middle Easterners. Each participant produced a written statement and participated in an investigative interview; word usage in both was analyzed. Word usage differentiated truths from lies in both the written statement and the investigative interview, and the effect sizes associated with these findings were substantial. For the written statement, word usage predicted truths from lies at 68.90% classification accuracy; for the investigative interview, word usage predicted truths from lies at 71.10% accuracy. Ethnicity did not moderate these effects. These findings are discussed in terms of their implications to cross-cultural applicability of the psychological demands placed on liars and in terms of their practical field utility. Copyright © 2014 John Wiley & Sons, Ltd.

Key words: veracity; lying; ethnicity; culture; LIWC

Several decades of research have provided rich sources of information concerning the behavioral indicators of veracity and lying, and this information has undoubtedly made major contributions to law enforcement, national security and intelligence practitioners. One source of indicators is verbal and comes from analyses of words used to provide statements and answer questions. Lies contain fewer words and omissions of information; are less plausible, structured and logical; are more internally discrepant and ambivalent; contain repeated details and lack contextual embedding; and include more descriptions of what did not occur (DePaulo *et al.*, 2003; Duran, Hall, McCarthy, & McNamara, 2010;

*Correspondence to: David Matsumoto, Department of Psychology, San Francisco State University, 1600 Holloway Avenue, San Francisco, CA 94132, USA.
E-mail: dm@sfsu.edu

Newman, Pennebaker, Berry, & Richards, 2003; Porter, Birt, Yuille, & Lehman, 2000; Porter & ten Brinke, 2010; Vrij, 2007).

There are two major methods of analyzing words for indicators of veracity and lying. One involves the human analysis of the semantic and grammatical structures associated with word usage involving a technique commonly known as Statement Analysis (SA). SA has its roots in psycholinguistic research in the early 1900s and its more modern roots in the work of Undeutsch (1989) and a technique known as Statement Validity Analysis (SVA). SVA was founded on a hypothesis that statements based on actual memories differ from statements based on fabrication or fantasy (Undeutsch, 1989). The crucial parts of SVA involve a criteria-based content analysis (CBCA) and an evaluation of CBCA outcomes using a Validity Check-List with criteria organized around categories such as general, unusual, motivational and stylistic features. In addition to SVA, a number of other techniques also exist, including Theoretical Verbal Analysis (TVA; Connelly *et al.*, 2006), Reality Monitoring (RM; Johnson & Raye, 1998), Scientific Content Analysis (SCAN; Sapir, 1996) and Investigative Discourse Analysis (IDA), which is an extension of CBCA and similar to SCAN (Rabon, 1994). These techniques are also based on principles concerning the nature of human memory and verbal recall.

A second method for analyzing words as indicators of veracity and lying involves identifying the words used and classifying them according to their psychological meaning and function without relying on their linguistic features or grammatical structures. Various techniques to do so exist, such as Coh-Metrix (Graesser, McNamara, Louwerse, & Cai, 2004), Leximancer (Smith & Humphreys, 2006) or the Linguistic Inquiry and Word Count (LIWC; Pennebaker, Francis, & Booth, 2001). Of these, LIWC is arguably one of the most commonly used and well-validated analytic tools in the literature of computerized linguistic analysis. Its output has been linked to a number of meaningful, underlying psychological dimensions (Tausczik & Pennebaker, 2010), and it has been used to examine differences in text between truths and lies. This paper focuses on the analysis of word usage.

A number of studies have documented word usage differences between truth telling and lying using LIWC. Across five studies, Newman *et al.* (2003) demonstrated that liars used less words overall, words with less cognitive complexity, more negative emotion, and fewer self-referent and other referents. Bond and Lee (2005) demonstrated that truthful statements of incarcerated prisoners included more sensory and perceptual processes when they told the truth compared to when they lied. Other studies have reported similar results (Duran, Hall, McCarthy, & McNamara, 2010; Hancock, Curry, Goorha, & Woodworth, 2008; Schelleman-Offermans & Merckelbach, 2010; Toma & Hancock, 2012; Vrij, Mann, Kristen, & Fisher, 2007). To be sure, findings have not always been consistent; two studies reported negative findings on words related to complexity and pronoun use (Hancock *et al.*, 2008; Toma & Hancock, 2012). A meta-analysis of studies using LIWC to distinguish truthful and deceptive statements, however, reported that deceptive statements tend to have more words expressing both positive and negative emotions, motion verbs and negation, and less self-referents and other referents, exclusive words, words related to perceptual processes, tentative words and time-related words (Hauch, Blandon-Gitlin, Masip, & Sporer, 2012).

One very important aspect to consider when understanding this literature concerns the context within which truths and lies were examined. Researchers have utilized a wide variety of contexts. One aspect of context is the type of truths and lies studied. For example, Newman and colleagues (2003) examined word usage related to abortion attitudes, likes and dislikes of others, and mock crime. Bond and Lee (2005) examined the

statements of prisoners when they either told the truth or lied about neutral or crime-related videos they watched. Hancock and colleagues (2008) and Duran and colleagues (2010) examined truths and lies told among unacquainted individuals in an initial, icebreaker interaction. Toma and Hancock (2012) studied truths and lies in online dating profiles, while Schelleman-Offermans and Merckelbach (2010) asked participants to write true and false statements about an aversive situation in which they were the victim.

Another aspect of context concerns the motivations of the truth tellers and liars, which is related to perceived stakes associated with their performances. Previous research has generally examined word usage produced in low-stake situations in which participants were not very motivated one way or another to lie or tell the truth, and there were little or no consequences if they were believed or not. Higher-stakes situations are those in which individuals are motivated to lie or tell the truth because there are perceived consequences if they are believed or not. Higher-stakes studies are more analogous to the real-life situations that face law enforcement and security personnel and are arguably more relevant to investigative psychology. Importantly, the behavioral indicators associated with veracity and lying that emerge from higher-stakes studies are different and more compelling than those that emerge from other studies (DePaulo *et al.*, 2003; Frank & Svetieva, 2013). Identifying indicators that are based in low-stakes studies that are not analogous to real-life situations and then training law enforcement personnel on them has dire consequences; at least one study demonstrated detrimental effects of training to detect lies when nonvalidated indicators are used (Kassin & Fong, 1999).

Another aspect of context concerns the source of the words analyzed. Some studies have analyzed written statements, others oral statements produced in interviews or individual declarations. Verbal markers of deception may differ depending on the type of source analyzed. When writing a statement, individuals have some time to collect their thoughts, reflect on the events to be recalled and choose their words to some degree. In a live interview, however, individuals are put on the spot as they interact with another person, listening and processing questions in relation to their lies while maintaining their composure. Thus, spoken words may differ than written words, resulting in differences in the linguistic markers of truth telling and lying.

A final aspect of context involves participant characteristics, and one of these is ethnicity. Little is known about whether or not the same linguistic markers of veracity and lying are equally valid when people of different ethnic or cultural backgrounds whose first language is not English produce statements because most of the research to date has analyzed source materials produced by native English speakers. To be sure, there are studies examining verbal indicators of veracity and lying in non-English languages (Blandon-Gitlin, Pezdek, Lindsay, & Hagen, 2009; Masip, Bethencourt, Lucas, Sanchez-San Segundo, & Herrero, 2012; Ruby & Brigham, 1997; Schelleman-Offermans & Merckelbach, 2010; Vrij, Akehurst, Soukara, & Bull, 2004). On the one hand, it is difficult to compare results across studies because of study differences (such as different procedures to elicit lies and different types of lies), which confound differences in results across studies. On the other hand, study differences also speak to the robustness of the findings. And only a handful has examined word usage.

Evaluating research on word usage differences between truths and lies according to these aspects of context leads to interesting insights about the findings to date. While the studies reviewed above have undoubtedly contributed much to an understanding of the words used by truth tellers and liars and the possible psychological mechanisms underlying word choice when lying, they may have limited potential applicability to actual investigative

contexts, especially involving lies about criminal acts. Investigative interviews often occur with suspects after a crime, are associated with high stakes if one is not believed (even if one tells the truth), can involve written statements (before or after the interview) as well as oral statements during an interview, and can involve people of different ethnic or cultural backgrounds. The available literature on word usage and deception, however, includes only one study that examined lies after a mock crime (Newman *et al.*, 2003, Study 5); all studies involved low-stakes situations; studies generally examined either written statements or words produced during very brief oral statements or questioning but not both, and no study has examined word usage differences in people of different ethnic/cultural backgrounds.

We addressed this gap by examining differences in word usage between truth tellers and liars in a moderately high stakes, real-life scenario and involving participants from four cultural/ethnic groups—European-Americans, Chinese, Hispanics and Middle Easterners. They participated in an adapted version of a mock-theft scenario in which they were either asked to steal a check made out to ‘Cash’ and lie about it or not steal a check and tell the truth. The Chinese, Hispanic and Middle Eastern groups were chosen to sample a broad range of cultural differences that may impact indicators of veracity. After being assigned to the steal–lie/do not steal–truth condition, each participant engaged in three interviews, two prior to committing the crime (screening and secondary interviews) and one afterwards (investigative interview). Prior to the third investigative interview, participants were asked to write a statement about everything they did while in the room where they could steal the check. Based on previous research (reviewed earlier), we identified specific LIWC variables that differentiated truths from lies and tested whether they did so in the written statements and one of the interviews obtained in this study; we also examined whether ethnicity moderated these effects. The specific words tested were those related to the following linguistic markers: total words used, cognitive complexity, self-referent and other referents, sensory and perceptual processes, positive and negative emotions, motion verbs, negation words, exclusive words, tentative words, and time-related words. We hypothesized that the words used in the written statements prior to the investigative interview and in the interview itself would differentiate truths from lies and that ethnicity would moderate those effects.

METHODS

Participants

Participants came from one of four ethnic/cultural groups: European Americans and Chinese, Hispanic and Middle Eastern immigrants. All participants were recruited from student and nonstudent communities in the San Francisco Bay Area and Buffalo, NY, through ads seeking ‘European American’, ‘Chinese’, ‘Hispanic’ or ‘Middle Eastern’ participants. The European Americans were all US born-and-raised Caucasians. Ads for the other three ethnic groups stipulated individuals who were either immigrants born and raised in their home country or first generation born in the US, whose first language was not English but that of the home country and whose parents were both born and raised in the home country. For the purposes of this study, home country was defined for Chinese as the People’s Republic of China, Hong Kong or Taiwan and first language was Mandarin or Cantonese; for Hispanics, country was defined as any country in Central or South America and first language was Spanish; and for Middle East, country was any country in Northern Africa and Western Asia and first language was Arabic. Prior to participation,

all potential participants were telephone screened according to the inclusion criteria recruited for and answered the same questions in a standard demographics form obtained as part of the pre-session measures (see below). Thus, the participants included in the study were those who self-identified as one of the ethnic groups, self-selected to contact the research team and confirmed their self-identification in screening and again as part of the pre-session measures. Additionally, all participants completed ratings of their ethnic identities (the General Ethnicity Questionnaire, below), which further confirmed the ethnic differences among the groups.

The inclusion criteria meant that the three immigrant samples consisted of participants who were descendants of home countries that were culturally different (e.g. China, Hong Kong and Taiwan). (To be sure, the same could be said about European Americans.) Ethnicity refers to people of a nation or tribe and can denote one's racial, national or cultural origins (Matsumoto & Juang, 2013). For example, within the US, African Americans, Asians and Pacific Islanders, Hispanics and Latinos, and Native Americans are often considered different ethnic groups. Culture refers to a unique meaning and information system, shared by a group and transmitted across generations, that allows the group to meet basic needs of survival and pursue well-being (Matsumoto & Juang, 2013), and ethnic groups are often markers of cultural differences. Cultural values data examining differences within a country and country differences within a world region, for instance, demonstrate that differences within a world region are smaller than differences between regions (Hofstede, 2001; Hofstede, Garibaldi de Hilall, Malvezzi, Tanure, & Vinken, 2010; Schwartz, 2004). Thus, we were fairly confident that the regional origins of the three immigrant groups represented meaningful cultural differences.

The inclusion criteria used also meant that some individuals in the immigrant groups were foreign nationals while others were technically US citizens. This is one of the reasons why we included the General Ethnicity Questionnaire in order to document group differences in it. The ethnic groups sampled represented the same ethnic group categories with which differences in expressivity and cultural norms for expression have been documented within the US (Matsumoto, 1993; Tsai & Levenson, 1997; Tsai, Levenson & Carstensen, 2000a). The ethnic groups in this study were also representative of the cultural and ethnic diversity that law enforcement officers in the US and other multicultural societies face.

The final sample included 226 individuals who participated for a cash payment (standard participation fee was \$20, with the possibility of making more depending on outcomes described below). The European Americans included $n_s = 40$ and 38 in the lie and truth conditions, respectively; the Chinese, Hispanics and Middle Eastern samples included $n_s = 46$ and 36, $n_s = 28$ and 18, and $n_s = 8$ and 12, respectively. Seventy percent of the entire sample was comprised of students (none of whom were psychology majors who may have participated in a similar experiment), and they were roughly evenly distributed between males (47.4%) and females (52.6%) with an average age of 27.32 ($SD = 10.42$, range 18–72). Sex and age breakdown were roughly equivalent across the four ethnic groups and within conditions. All experimental procedures occurred in English.

Measures

At the beginning of the experiment, all participants completed a basic demographics questionnaire, the General Ethnicity Questionnaire (GEQ; Tsai, Ying, & Lee, 2000b), the Machiavellianism Scale (Christie, 1970) and the Self-Monitoring Scale (Snyder, 1974). Participants also completed an emotion checklist at the beginning and end of the

experiment. This checklist included 12 emotion words (guilt, fear, anger, embarrassment, worry, contempt, excitement, disgust, amusement, nervousness, surprise and interest) rated on nine-point scales labeled 0 = *None*, 4 = *Moderate Amount* and 8 = *Extremely Strong*.

The GEQ is a commonly used scale to measure acculturation and ethnic identity and was included as a manipulation check for ethnic/cultural differences. This questionnaire contains 38 statements, 25 rated on a five-point Likert scale from strongly disagree to strongly agree and 13 rated on a five-point scale from very much to not at all. The GEQ was modified to be applicable to each ethnic group. Analyses of the GEQ Total score, which was the mean of all items after reverse coding those negatively loaded, indicated that our Chinese sample had significantly higher scores than American born Chinese and Chinese who immigrated to the US before the age of 12 reported by Tsai *et al.* (2000b), $t(74) = 8.07$, $p < .001$, $d = .93$; $t(74) = 1.71$, $p < .05$, $d = .20$, respectively. These analyses demonstrated that our Chinese sample identified themselves as Chinese and strongly with Chinese culture more so than American born Chinese. Norms for Hispanics and Middle Easterners using this same measure do not exist, but their scores were comparable to the Chinese in our sample.

Interviewers and questions

Ten male actors, all above the age of 30, served as interviewers. The authors, who have experience training law enforcement officers in investigative interviewing, conducted training sessions for the interviewers, some of whom were former law enforcement officers, to deliver the interviews in a neutral and objective manner and to stick with the predetermined interview questions. The interviewers worked on a rotating schedule such that three were present for each session. Each interviewer's specific role alternated daily between Interviewer 1 (I1), Interviewer 2 (I2) or Interviewer 3 (I3); the first author also served as an interviewer.

Of the 10 interviewers, five were Caucasians, three were African Americans, one was Middle Eastern and one was Asian. They ranged in age from 35 to 52. Ethnicity was not kept constant across the interviewees; interviewers were scheduled according to their availability, once training had occurred and it was determined that they could deliver the interviews in a fairly uniform manner.

The questions used in all three interviews were modeled after questions used in real-life security and investigative interviewing situations. They were developed after consultation with Subject Matter Experts (SMEs) from various law enforcement entities with interests in the practical application of the findings. Thus, the questions were designed to be as realistic as possible yet to retain fidelity for research purposes. The questions for the first two interviews were based on those typically used at checkpoint security situations. For the postevent investigative interview, we incorporated questions typically used by law enforcement officers (e.g. bait and indicator questions) as well as unanticipated questions and questions based on the Strategic Use of Evidence (SUE) technique (Hartwig, Granhag, Stromwall, & Kronkvist, 2006; Hartwig, Granhag, Stromwall, & Vrij, 2005). The initial screening included seven questions and lasted an average of 1:56 m; the secondary screening included 14 questions and lasted an average of 4:55 m. The investigative interview included a total of 18 questions and lasted an average of 9:46 m; of these, six were selected for analysis in this study because they were questions to which liars had to commit themselves directly to their lies or were questions that could theoretically differentiate truth tellers from liars (bait and indicator questions):

- Tell me in as much detail as possible what you did right after you left the previous interview.
- Describe in detail everything you did in the file room.
- I know that the check was made out to cash. It is very understandable that you would want to earn additional money by taking that check. If I were you, I would have wanted to take it. Now tell me truthfully, did you take the check?
- Why do you think that someone would take this \$100 check?
- What should happen to someone who steals money and is caught?
- I have evidence of you with the white envelope. We have a security camera that takes snapshots every 30 s, and we have one of you holding the white envelope. What do you have to say about that?

Stakes

Participants were told they will earn a minimum of \$20 for their participation and bonuses of either \$0 or \$80 depending upon their assigned condition and the determinations of the interviewers. If they took the check and were believed by all three interviewers, they received an additional \$80 and were allowed to leave early; but if they took the check and were not believed by any one interviewer, they received no additional money and had to stay an additional hour filling out a long questionnaire. If they did not take the check and were believed by all three interviewers, they received an additional \$10 and were allowed to leave early; but if they did not take the check and was not believed by any one interviewer, they received no additional money and had to stay an additional hour filling out a long questionnaire.

The stakes were different when lying and telling the truth because they reflected the stakes that occur in real life for the type of investigative interview examined in this study. Being a successful liar is likely associated with large rewards in real life; and participants who may not be inclined to do so in the first place require additional motivation to do so. As it is easier for people to tell the truth, there are indeed rewards for telling the truth, but they are lower than when successfully lying. If, however, the type of investigative interview were different, the stakes applied in this study may not be the most appropriate. For example, a context in which a truth teller were falsely accused of committing a crime would probably require a larger stake in being perceived as truthful. This, however, would require a different study. Thus, readers are cautioned to interpret the findings reported below vis-à-vis the particular way in which the experiment was conducted, including the stakes involved.

As a manipulation check, participants were asked to rate the severity of the above consequences if they were judged to be lying in the experiment, using a scale from 1 = *No consequence, even slightly pleasurable* to 10 = *Maximum consequence, even slightly painful*. The rating was done after the instructions and stakes were explained to the participants, and after the participants were assigned their condition and reminded of their tasks and stakes.

Procedures

Participants were introduced to the study and told that they would be randomly assigned to either take a \$100 check made out to cash or to look at but not take the \$100 check. They were also told that irrespective of the assignment, their goal was to go through up to three

checkpoints/interviews, convincing the officers of their honesty and sincerity. In order to be judged as truthful, each participant needed to convince all officers of their innocence. The stakes associated with the experiment were revealed.

After instructions were delivered and consent obtained, participants completed the pre-session measures. When done, the experimenter conducted a random assignment to condition procedure by asking the participant to select one of 10 cards, five for the truth condition and five for the lie condition. Participants were given a folder containing their condition specific instructions, including a map of the layout of the area and the location of the room where the stealing of the check would take place. The instructions repeated the stakes associated with their successful or unsuccessful performance and where to find the check and what to do with it (either take it or look at it and put it back). All participants were instructed to take a red index card in order to receive experiment payment; this card also constituted a story that the participants could use to communicate to the interrogators why they were going to the file room. The participants were asked to keep the instruction reminder folder and to review the instructions later before going to the file room. The participants were thereafter given a quick verbal reminder about their instructions and the payoff stakes.

When the participant was ready to proceed, a second experimenter entered who was blind to the participant's condition assignment, as were all interviewers. This experimenter escorted the participant out of the instruction area to the main floor of the experimental area near a metal detector. The experimenter then left and Interviewer 1 (I1) appeared, walked by the participant, stepped behind a checkpoint interview table and called the participant to empty their pockets before going through the metal detector. I1 then conducted the screening interview.

When the screening interview was done, I1 left and the experimenter reentered the room and instructed the participant to have a seat in a waiting area. After a few minutes, the experimenter reentered and informed the participant that he or she had been selected for a second interview and should go to the interview room. In the interview room the participant was left alone to sit and wait. Interviewer 2 (I2) entered and conducted the secondary screening interview.

Upon completion, I2 exited the room and the experimenter reentered and escorted the participant back to the waiting area. The participant was instructed to reexamine the instruction reminder sheet and to perform his or her assigned task in the file room. The experimenter exited and when ready the participants went to the file room, clearly labeled. In the room, the participants found a folder that contained the check and either looked at the check or took it by putting it on his or her body. In addition, all participants were asked to take the red index card. Once ready, the participants returned to the waiting area and rang a bell. The experimenter reentered and informed the participant that he or she had been selected for a third interview and escorted the participant to the interview room.

Interviewer 3 (I3) entered the interview room and asked the participant to write a statement about 'everything that happened in the file room' and left the room to give the participant time to write the statement. After 3 min, I3 reentered the room, read the statement quickly and conducted the investigative interview. I3 left the interview room upon completion. The experimenter reentered and reclaimed the participant by escorting him or her back to the original instruction and consent room and then left. The first experimenter reentered. Post-session measures were administered along with the debriefing, post-session consent and payment.

The design of the experiment meant that lying was always associated with stealing the check, while truth telling was always associated with not stealing, introducing a potential

confound between lying/truth telling and stealing/not stealing, which may have influenced the word usage of the participants. In actuality, we could have crafted the study so that half the participants stole and half did not, and both halves either told the truth or lied about it. Although this design appears clean, however, in fact it introduces a number of other concerns. For example, if participants were asked to take the check but tell the truth about it, then technically they would not have stolen it, and thus, their verbal output would not be associated with 'stealing'. Also, if the participants were asked to not take the check but to lie about it, it would mean that they would have had to lie by saying that they took the check during the interviews, which does not make sense in relation to the questions asked. Thus, we opted to conduct the study with the two conditions that we did and suggest that it is closer to being ecologically valid than the introduction of the other two conditions, at least for the type of investigative context we utilized.

Linguistic analyses

We transcribed the written statements and participants' responses in the third investigative interview and used the LIWC to analyze the words used in those statements (the 2007 dictionary was used). Transcripts were prepared according to the LIWC manual and were corrected for errors prior to processing.

LIWC counts the number of words in a body of text that correspond to various categories of meaning and converts the tallies into percentages of the total text (Pennebaker *et al.*, 2001). It uses an internal dictionary comprised of several word categories to classify how much a group of words relate to a particular topic. This dictionary is composed of about 4500 words and word stems, each of which defines one or more nonmutually exclusive word categories in a hierarchical order (e.g. anger words are categorized as anger, negative emotion and overall emotion words). The LIWC word categories have adequate psychometric properties (Pennebaker *et al.*, 2001).

Each word in a source document is compared with words in the dictionary file and if a match occurs the appropriate category(ies) for that word is tallied; various structural composition elements (e.g. word count and sentence punctuation) are also counted. Output categories include general descriptors (total word count, words per sentence, percentage of words captured by the dictionary and percent of words longer than six letters), linguistic dimensions (e.g., pronouns, articles and auxiliary verbs), psychological constructs (e.g. affect, cognition and biological processes), personal concerns (e.g. work, home and leisure activities), paralinguistic dimensions (assents, fillers, nonfluencies) and punctuations (periods, commas, etc.) (For a more complete description of the LIWC processing procedures and its development, see Pennebaker, Chung, Ireland, Gonzales, & Booth, 2007).

As mentioned earlier, we selected for testing in this study words that were found to differentiate truth tellers and liars in previous research on word usage. The specific LIWC variables tested were as follows (LIWC variable names in parentheses):

- Total words used (Word count)
- Cognitive complexity (the sum of Excl and Negate, per Chung and Pennebaker, 2007; variables were standardized prior to summing)
- Self-referent and other referents (I, We, You, She, They)
- Sensory and perceptual processes (See, Hear, Feel)
- Positive and negative emotions (Posemo, Negemo)
- Motion verbs (Motion)

- Negation words (Negate)
- Exclusive words (Excl)
- Tentative words (Tentat)
- Time-related words (Time)

Coding of the investigative interview

Words that are produced in any interview are highly dependent on many factors. Because we wanted to isolate the effects of the conditions on word usage and eliminate potential confounds of issues related to the interview, we coded the interviews in two ways. First, we identified statements when it was apparent that a participant did not understand the relevant question being posed (coded Yes or No). Examples included a participant asking the interviewer to repeat the question multiple times or a participant providing a response that clearly did not answer the question. Second, we identified instances when the interviewer impeded or negatively influenced the interview process, thereby potentially causing the participant to provide inaccurate information. The following examples of interviewer contamination occurred: the interviewer misstated or rearranged the words of the relevant question so that it altered the meaning of the original question, interrupted a participant when he or she was responding, interjected words or phrases during a participant's response such as 'keep going' and 'go on' or volunteered words to help a participant complete a response. Specific questions for which interviewer contamination occurred were identified (coded Yes or No).

Two coders, both of whom had several decades of law enforcement experience and who had extensive experience in analyzing word usage in real-life investigative settings, independently coded transcripts from 30 cases. Both coders were blind to the condition assignment of all cases. Reliabilities were high and acceptable for both Participant did not Understand and Interviewer Contamination codes ($rs = .97$ and $.83$, respectively). A single coder then coded the remainder of the cases.

RESULTS

Manipulation checks

Several analyses served as manipulation checks. First, we examined the participants' ratings of the perceived consequences of the stakes described in the experiment. The mean rating was slightly above the midpoint, $M = 5.68$, $SD = 2.24$, suggesting that the stakes were perceived to be of moderate severity to the participants. There were no ethnicity or condition differences in these consequence ratings.

We also examined presession to postsession changes in self-reported emotions. We computed a Prepost (2) \times Emotion (12) \times Veracity Condition (2) mixed ANOVA. The three-way interaction was significant, $F(11, 2046) = 4.96$, $p = .000$, $\eta_p^2 = .03$, as was the Prepost \times Veracity Condition interaction, $F(1, 186) = 18.49$, $p = .000$, $\eta_p^2 = .09$. To decompose the significant three way, we computed single-*df*, Prepost \times Veracity Condition simple interaction contrasts separately for each of the emotions. Significant effects were found for seven emotions $.02 < \eta_p^2 < .19$. Participants who stole the check and lied about it reported significant increases in guilt, fear, embarrassment, worry and nervousness compared to participants who told the truth. Participants who did not steal the check and told the truth

about it reported significantly less excitement and interest than participants who lied. (Tables of means and analyses are available from the authors.) These analyses suggested that the participants perceived the stakes to be of moderate severity and that they were emotional about their participation.

Main analyses

In order to examine whether word usage differed as a function of veracity condition and participant ethnicity, we computed an overall Ethnicity (4) by Veracity Condition (2) MANOVA on the selected LIWC variables separately for the written statement and the investigative interview. For the written statements, the main effect of Veracity Condition was significant, $\lambda = .72$, $F(16, 202) = 4.34$, $p < .001$, $\eta_p^2 = .26$. Neither the main effect of Ethnicity nor the interaction was significant, $\lambda = .78$, $F(48, 601.59) = 1.25$, ns, $\eta_p^2 = .08$ and $\lambda = .77$, $F(48, 601.59) = 1.14$, ns, $\eta_p^2 = .08$, respectively. Consistent with the recommendations of Huberty and Morris (1989), we followed significant main effects with binary logistic regressions using Veracity Condition as the dependent variable and the LIWC variables as covariates. [Actually, Huberty and Morris (1989) recommended the use of discriminant analysis after significant MANOVA main effects. We opt for log regressions as they accomplish the same task as discriminant analysis but with better estimates of error.] Because the LIWC variables constituted a pool of variables that previous research had identified as differentiating truths from lies, we utilized backward conditional exclusion criteria (use of forward conditional inclusion criteria produced essentially the same results). The final equation included seven LIWC variables—Word count, See, Posemo, Negate, Tentat, Motion and Time—and accounted for 68.90 overall correct classification of cases (Table 1, top).

For the investigative interview, the main effect of Veracity Condition was marginally significant, $\lambda = .89$, $F(14, 189) = 1.72$, $p < .06$, $\eta_p^2 = .11$. When we recomputed the overall MANOVA by filtering the data to include only participants whose interviews had no evidence of interviewer contamination or participant misunderstanding, however, the main effect of Veracity Condition was significant, $\lambda = .73$, $F(16, 98) = 2.29$, $p < .01$, $\eta_p^2 = .27$ (the filtering process resulted in a total $N = 121 - n_{\text{Chinese}} = 44$, $n_{\text{European American}} = 29$, $n_{\text{Hispanic}} = 35$ and $n_{\text{Middle Eastern}} = 13$). We then computed logistic regressions using Veracity Condition as

Table 1. Final results of logistic regressions, separately for written statements and investigative interviews

Final Model Chi-Square	Overall Correct Classification (%)	False Positive (%)	False Negative (%)	Variables In	B	SE
Written Statement						
$\chi^2(7, 225) = 58.37$, $p < .001$	68.90	14.67	16.44	See	-.08	.04
				Posemo	.22	.12
				Negate	-.51	.15
				Tentat	.52	.15
				Time	-.14	.06
				Word count	-.03	.01
				Motion	-.13	.06
Investigative Interview						
$\chi^2(3, 121) = 24.17$, $p < .001$	71.10%	19.01%	9.92%	I	.16	.08
				Shehe	1.14	.55
				Motion	-.57	.15

the dependent variable and the LIWC variables as covariates on the filtered data, using backward conditional exclusion criteria. The final equation included three LIWC variables—I, Shehe and Motion—and accounted for 71.1% overall correct classification of cases (Table 1, bottom).

To examine whether the same pattern of results occurred separately for each of the four ethnic groups, we examined the means for each of the significant LIWC variables listed in Table 1. For the written statement, the same pattern of results occurred for all four ethnicities for Word count, See, Posemo and Tentat, and for three of the four groups for Time and Motion. For the investigative interview, the same pattern of results occurred for all four ethnicities for I and Motion, and for three of the four groups for Shehe. Detailed table of means is available from the authors upon request.

Posthoc analyses

Veracity condition differences

To investigate other possible word usage differences that could differentiate truths from lies, we computed posthoc, one-way ANOVAs on the nonselected LIWC variables using Veracity Condition as the factor. Because of the large number of significance tests, we adopted a $p < .01$ criterion with which to evaluate statistical significance. For the written statement, truths included more articles and words related to money, while lies included more auxiliary verbs. For the investigative interviews, truths included more words related to relativity and money, while lies included more verbs, auxiliary verbs and words related to sadness and insight (Table 2).

Ethnicity differences

The main effect of Ethnicity in the overall MANOVA using the filtered investigative interview data was significant, $\lambda = .53$, $F(42, 297.41) = 1.69$, $p < .01$, $\eta_p^2 = .19$. We followed this

Table 2. Results of posthoc analyses on nonselected LIWC words

Variable	Truths	Lies	<i>F</i>
Written Statement			
Articles	14.13 (4.30)	12.59 (4.95)	$F(1, 217) = 13.48$, $p < .001$, $\eta_p^2 = .06$
Auxiliary verbs	3.41 (3.45)	4.55 (3.76)	$F(1, 217) = 7.86$, $p < .005$, $\eta_p^2 = .04$
Money	2.46 (1.78)	1.08 (183)	$F(1, 217) = 15.69$, $p < .001$, $\eta_p^2 = .07$
Investigative Interview			
Verb	16.28 (3.35)	17.80 (3.68)	$F(1, 113) = 7.30$, $p < .01$, $\eta_p^2 = .06$
Auxiliary verbs	7.61 (2.75)	8.88 (2.82)	$F(1, 113) = 10.21$, $p < .01$, $\eta_p^2 = .08$
Sadness	.06 (.16)	.10 (.21)	$F(1, 113) = 6.94$, $p < .01$, $\eta_p^2 = .06$
Insight	2.00 (1.34)	2.80 (1.79)	$F(1, 113) = 7.53$, $p < .01$, $\eta_p^2 = .06$
Relative	16.52 (4.50)	14.43 (4.39)	$F(1, 113) = 10.00$, $p < .01$, $\eta_p^2 = .08$
Money	2.45 (1.21)	1.68 (1.15)	$F(1, 113) = 10.16$, $p < .01$, $\eta_p^2 = .08$

effect by examining univariate ethnicity differences separately for each LIWC variable. The univariate effects for Shehe and Excl were significant, $F(3, 113) = 3.88$, $p < .05$, $\eta_p^2 = .09$ and $F(3, 113) = 4.46$, $p < .01$, $\eta_p^2 = .11$. Pairwise comparisons using Bonferroni corrections indicated that the Chinese had significantly higher means on Shehe than all other ethnic groups and that Middle Easterners had significantly higher means on Excl than did European Americans and Hispanics.

DISCUSSION

The results supported the hypothesis that the selected words would differentiate truths from lies in both the written statement and the investigative interview, and the effect sizes associated with the findings were substantial (η_p^2 s = .26 and .27, respectively). For the written statement, the selected LIWC variables significantly predicted truths from lies at 68.90% classification accuracy, with modest false positive and false negative classifications. For the investigative interview, the selected LIWC variables significantly predicted truths from lies at 71.10% classification accuracy, with a fairly low false negative rate. These classification rates were impressive given that the average correct classification rates for humans is about 54% (Bond & DePaulo, 2006). Ethnicity did not moderate these effects.

These findings were not generated without limitations. First, although the Chinese, Hispanic, and Middle Eastern participants were either first or second generation immigrants to the US, they all wrote and spoke in English. It was possible, therefore, that the nonfindings concerning ethnic differences occurred because the participants used English and use of a standard language diluted the possibility of finding ethnic differences. Literature suggesting code or cultural frame switching (Hong, Morris, Chiu, & Benet-Martinez, 2000) among bilinguals certainly is supportive of such a possibility. And, the multilingual/multicultural participants may have interpreted differently the study procedures, especially the stakes, and these differences may have influenced the data. Despite this possibility, we included bilinguals in the study because the GEQ data indicated that they were culturally different than the European Americans. Inclusion of bilinguals in the study was ecologically valid as there are many non-native English speaking individuals in the US (and in many countries) who come in contact with law enforcement and must speak or write in English, raising questions about the possibility of ethnic differences in the verbal indicators of veracity and lying.

Another limitation of the study concerned the exact questions that participants responded to in the investigative interview. While these questions had ecological validity and were grounded in the research literature, the responses given were inextricably tied to the questions asked; thus, the findings generated for the investigative interview were limited to the questions posed. If different questions were delivered to the participants, different responses would have been given, thereby producing different findings. Readers are therefore cautioned to interpret the findings above for the investigative interview with this caveat, which may have contributed to the nonfindings with regard to complexity. This was not as much a factor for the written statement, although certainly different instructions for it would also produce different statements.

A third limitation had to do with the differences in sample sizes across the ethnicities and especially the smaller sample size of Middle Eastern participants; this was particularly true when the data from the investigative interview were filtered for interviewer contamination. Differences in the sample sizes made statistical comparisons among the ethnicities difficult. The effect sizes should enable readers to gauge the meaningfulness of the effects regardless

of the associated *p* values. Still readers need to take into account the relatively small sample sizes of the Hispanics and Middle Easterners when interpreting the findings related to ethnicity differences.

Despite these limitations, the findings produced interesting results that deserve attention. In the written statement, truth tellers used more words related to sensory processes, negations, motion and time. These findings are commensurate with the Reality Monitoring framework (Johnson, 1988; Johnson & Raye, 1981), which predicts that words related to time, space and sensory processes occur in accounts based on actually experienced events. That truth tellers used more words in general was consistent with previous research (Newman *et al.*, 2003).

That liars used more tentative words than did truth tellers was contrary to previous findings (see Hauch *et al.*, 2012). The use of more tentative words by liars suggests that they wrote with some equivocation. Equivocation words qualify statements, which allow liars to distance themselves from the act or content of lying by tempering the action about to be described or by discounting the message even before it is transmitted (Weintraub, 1989).

That liars used more words related to positive emotions than did truth tellers was difficult to interpret. On one hand, Reality Monitoring would suggest that truth tellers use more words related to affect and emotion because those are associated with actual experiences. On the other hand, some researchers have suggested that liars would use more negative words because of guilt or anxiety associated with the act of lying (Vrij, 2008). A previous meta-analysis indicated that liars tend to use more words expressing both positive and negative emotion (Hauch *et al.*, 2012). That the liars in our study used more words related to positive emotions was consistent with the results of that meta-analysis. Future research examining the target of those emotional words—that is, what the participants were emotional about—may shed light on this issue.

Liars used more words related to first (I) and third person (Shehe) references in the investigative interview. That liars used more third person pronouns is consistent with the notion that the liars may have distanced themselves from the events that occurred because those references deflect attention away from the self and focus attention on others. This effect was consistent with those reported in previous research (Hauch *et al.*, 2012; Newman *et al.*, 2003; Sapir, 1996). That liars used less words related to motion is also consistent with the Reality Monitoring framework, which suggests that such words would be associated with truthful accounts because they are part of one's actual experiences. That liars used more first person personal pronouns (I), however, was contrary to what has been found in previous studies. Our anecdotal review of the transcripts from the investigative interview suggested that liars may have produced more statements that were extraneous to the direct questions asked, thus producing more statements that started with 'I'. In fact, contradictory findings have been reported in the literature (see review above), and future studies may examine whether the type of statement being made may account for the contradictions.

The fact that different words differentiated truths and lies in the written statement and investigative interview speaks to notion that researchers (and practitioners) need to pay attention to the specific source in which words are being analyzed to make determinations of veracity or deception. When writing statements, individuals have some time to collect their thoughts, reflect on the events to be recalled and choose their words to some degree. But in an interview, individuals are put on the spot as they are interacting with another person and must listen and process questions in relation to their lies while maintaining their composure. It is natural to use pronouns as subjects of sentences, and decisions concerning pronoun usage must be made quickly when speaking. Thus, it makes sense that differences

between truths and lies in pronoun usage occurred during the interviews, whereas differences occurred in word usage related to other cognitive processes in the written statement.

That ethnicity did not moderate the findings above was also important (but recall the limitations concerning ethnicity and sample size mentioned above). Consistencies in findings across cultural and ethnic groups provide additional evidence for potential pancultural similarity in the underlying psychological effects of lying and similarity as reflected in the linguistic choices that mark those effects. This consistency is commensurate with many single-language studies of non-English languages (Blandon-Gitlin *et al.*, 2009; Masip *et al.*, 2012; Ruby & Brigham, 1997; Schelleman-Offermans & Merckelbach, 2010; Vrij *et al.*, 2004). This suggests a psychological mechanism of lying that may be cross-cultural and adds to a small but growing list of studies demonstrating cross-cultural similarities in other aspects of deception, including in the attitudes and beliefs about lying and in judgments of lies (Bond & Atoum, 2000; Bond, Omar, Mahmoud, & Bonser, 1990; The Global Deception Research Team, 2006). There are also important practical ramifications for understanding consistency in findings for practitioners, so that interviews with individuals from different cultures and language groups can occur with the confidence of cross-ethnicity applicability of the techniques of deception detection.

Posthoc analyses revealed other word usage differences between truth tellers and liars that were fairly consistent with previous theoretical frameworks. That truth tellers used more words related to relativity or money was consistent with Reality Monitoring (Johnson, 1988; Johnson & Raye, 1981). That liars used less words related to articles is consistent with their greater use of equivocation, described above, because articles function to provide specificity in language (Chung & Pennebaker, 2007). To be sure, these findings were not originally predicted and should be replicated in future studies, as they may be limited to the methods of this study.

Posthoc analyses also demonstrated some ethnic group differences in overall rates of some word usages. These findings may be indicative of different linguistic styles of communication across ethnicities, despite the fact that all used English. Future research will need to replicate and extend these findings, and further investigate the interesting possibility of ethnic differences in word usage as part of linguistic style differences in communication. Practically, investigators will need to be aware of ethnicity main effects in the very same variables that differentiated veracity condition (e.g. third person pronouns) so that they can make more accurate interpretations of those markers when they occur in interviews and statements.

The findings overall and especially from the log regressions were indicative of the power of computerized linguistic analysis tools such as the LIWC to be able to differentiate truths from lies verbally. The 68.90% and 71.10% correct classification rates for the written statement and investigative interview were impressive given that the average accuracy rates for humans is about 54% (C. F. Bond & DePaulo, 2006). At the same time, it is important to remember that the findings reflected differences in the relative amounts of the respective variables between truths from lies, and the means for some variables were near zero (e.g. Sad in the posthoc analyses). Although such differences are important and interesting for research purposes, they raise questions about the field applicability of the findings for investigators. Investigators typically do not have a true and false statement or interview of the same person to compare to each other; they have a single statement or interview that they do not know is true or false or more precisely what parts are true and what parts are false. In the field, investigators do not have the luxury to compare relative amounts of different linguistic markers across different verbal samples

concerning the same event. Thus, linguistic analyses of verbal markers based on simple word counts may have limited application in the field, at least related to how such analyses are typically conducted and presented (as we have done here). Future research may examine how different algorithms or scoring methodologies may be produced that has more field utility.

Future research will also need to examine the verbal statements of people of different ethnicities and cultures in their native languages, not just English, and in relation to different types of crimes or transgressions. Whether the same psychological demands are placed on individuals in the same ways when speaking and writing in their native language is an open question, more importantly whether the same linguistic markers manifest those psychological demands is also a question ripe for future research. If the rules of grammar and deep structure of language (Chomsky, 1957, 1972) and the principles of memory and recall (Undeutsch, 1989) are similar across cultures, then verbal indicators of truths and lies should occur regardless of culture, ethnicity, and language.

ACKNOWLEDGEMENTS

This work was funded by the High-Value Detainee Interrogation Group contract J-FBI-12-197 awarded to Humintell LLC. Statements of fact, opinion and analysis in the paper are those of the authors and do not reflect the official policy or position of the FBI or the U. S. Government. The authors wish to thank Vincent A. Sandoval and Lisa G. Skinner for their aid as coders.

REFERENCES

- Blandon-Gitlin, I., Pezdek, K., Lindsay, D. S., & Hagen, L. (2009). Criteria-based content analysis of true and suggested accounts of events. *Applied Cognitive Psychology, 23*, 901–917.
- Bond, C. F., & Atoum, A. O. (2000). International deception. *Personality and Social Psychology Bulletin, 26*, 385–395.
- Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10*, 214–234.
- Bond, C. F., Omar, A., Mahmoud, A., & Bonser, R. N. (1990). Lie detection across cultures. *Journal of Nonverbal Behavior, 14*, 189–204.
- Bond, G. D., & Lee, A. Y. (2005). Language of lies in prison: Linguistic classification of prisoners' truthful and deceptive natural language. *Applied Cognitive Psychology, 19*, 313–329.
- Chomsky, N. (1957). Syntactic structures. New York: Walter de Gruyter.
- Chomsky, N. (1972). Language and mind. New York: Harcourt, Brace, and Johanovich.
- Christie, R. (1970). Scale construction. In R. Christie, & F. L. Geis (Eds.), *Studies in Machiavellianism* (pp. 10–34). New York: Academic Press.
- Chung, C., & Pennebaker, J. (2007). The psychological functions of function words. In K. Fiedler (Ed.), *Social communication* (pp. 343–359). New York: Psychology Press.
- Connelly, S., Allen, M. T., Ruark, G. A., Kligyte, V., Waples, E. P., Leritz, L. E., & Mumford, M. D. (2006). Exploring content coding procedures for assessing truth and deception in verbal statements, Year 3, Cumulative Final Report. Norman, OK: University of Oklahoma.
- DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129*, 74–118.
- Duran, N. D., Hall, C., McCarthy, P. M., & McNamara, D. S. (2010). The linguistic correlates of conversational deception: Comparing natural language processing technologies. *Applied Psycholinguistics, 31*, 439–462.

- Frank, M. G., & Svetieva, E. (2013). Deception. In D. Matsumoto, M. G. Frank, & H. S. Hwang (Eds.), *Nonverbal communication: Science and applications* (pp. 121–144). Newbury Park, CA: Sage.
- Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments and Computers*, *36*, 193–202.
- Hancock, J. T., Curry, L. E., Goorha, S., & Woodworth, M. (2008). On lying and being lied to: A linguistic analysis of deception in computer-mediated communication. *Discourse Processes*, *45*, 1–23.
- Hartwig, M., Granhag, P. A., Stromwall, L. A., & Kronkvist, O. (2006). Strategic use of evidence during police interviews: When training to detect deception works. *Law and Human Behavior*, *30*, 603–619.
- Hartwig, M., Granhag, P. A., Stromwall, L. A., & Vrij, A. (2005). Detecting deception via strategic disclosure of evidence. *Law and Human Behavior*, *29*, 469–484.
- Hauch, V., Blandon-Gitlin, I., Masip, J., & Sporer, S. L. (2012, 23–27 April 2012). Linguistic cues to deception assessed by computer programs: A meta-analysis. Paper presented at the Proceedings of the EACL 2012 Workshop on Computational Approaches to Deception Detection, Avignon, France.
- Hofstede, G. H. (2001). *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations* (2nd ed). Thousand Oaks, CA: Sage Publications.
- Hofstede, G. H., Garibaldi de Hilall, A. V., Malvezzi, S., Tanure, B., & Vinken, H. (2010). Comparing regional cultures within a country: Lessons from Brazil. *Journal of Cross-Cultural Psychology*, *41*, 336–352.
- Hong, Y.-y., Morris, M., Chiu, C.-Y., & Benet-Martinez, V. (2000). Multicultural minds: A dynamic constructivist approach to culture and cognition. *American Psychologist*, *55*, 709–720.
- Huberty, C. J., & Morris, J. D. (1989). Multivariate analysis versus multiple univariate analyses. *Psychological Bulletin*, *105*, 302–308.
- Johnson, M. K. (1988). Reality monitoring: An experimental phenomenological approach. *Journal of Experimental Psychology: General*, *117*, 390–394.
- Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review*, *88*, 67–85.
- Johnson, M. K., & Raye, C. L. (1998). False memories and confabulation. *Trends in Cognitive Sciences*, *2*, 137–145.
- Kassin, S. M., & Fong, C. T. (1999). 'I'm innocent!': Effects of training on judgments of truth and deception in the interrogation room. *Law and Human Behavior*, *23*, 499–516.
- Masip, J., Bethencourt, M., Lucas, G., Sanchez-San Segundo, M., & Herrero, C. (2012). Deception detection from written accounts. *Scandinavian Journal of Psychology*, *53*, 103–111.
- Matsumoto, D. (1993). Ethnic differences in affect intensity, emotion judgments, display rule attitudes, and self-reported emotional expression in an American sample. *Motivation & Emotion*, *17*, 107–123.
- Matsumoto, D., & Juang, L. (2013). *Culture and psychology* (5th ed). Belmont, CA: Cengage.
- 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*, 665–675.
- Pennebaker, J. W., Chung, C. K., Ireland, M., Gonzales, A., & Booth, R. J. (2007). The development and psychometric properties of LIWC2007. Austin, TX: LIWC.net.
- Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). *Linguistic inquiry and word count*. Mahwah, NJ: Lawrence Erlbaum.
- Porter, S., Birt, A. R., Yuille, J. C., & Lehman, D. R. (2000). The negotiation of false memories: The influence of interviewer and rememberer characteristics on memory distortion. *Psychological Science*, *11*, 513–516.
- Porter, S., & ten Brinke, L. (2010). The truth about lies: What words in detecting high-stakes deception? *Legal and Criminological Psychology*, *15*, 57–75.
- Rabon, D. (1994). *Investigative discourse analysis*. Durham, NC: Carolina Academic Press.
- Ruby, C. L., & Brigham, J. C. (1997). The usefulness of the criteria-based content analysis technique in distinguishing between truthful and fabricated allegations: A critical review. *Psychology, Public Policy, and Law*, *3*, 705–737.
- Sapir, A. (1996). *The laboratory for scientific interrogation course on scientific content analysis (SCAN) workbook*. Phoenix, AZ: Laboratory for Scientific Interrogation.
- Schelleman-Offermans, K., & Merckelbach, H. (2010). Fantasy proneness as a confounder of verbal lie detection tools. *Journal of Investigative Psychology and Offender Profiling*, *7*, 247–260.

- Schwartz, S. H. (2004). Mapping and interpreting cultural differences around the world. In H. Vinken, J. Soeters, & P. Ester (Eds.), *Comparing cultures, dimensions of culture in a comparative perspective* (pp. 43–73). Leiden, The Netherlands: Brill.
- Smith, A. E., & Humphreys, M. S. (2006). Evaluation of unsupervised semantic mapping of natural language with Leximancer concept mapping. *Behavior Research Methods*, *38*, 262–279.
- Snyder, M. (1974). Self-monitoring of expressive behavior. *Journal of Personality and Social Psychology*, *30*, 526–537.
- Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, *29*, 24–54.
- The Global Deception Research Team. (2006). A world of lies. *Journal of Cross-Cultural Psychology*, *37*, 60–74.
- Toma, C. L., & Hancock, J. T. (2012). What lies beneath: The linguistic traces of deception in online dating profiles. *Journal of Communication*, *62*, 78–97.
- Tsai, J. L., & Levenson, R. W. (1997). Cultural influences of emotional responding: Chinese American and European American dating couples during interpersonal conflict. *Journal of Cross-Cultural Psychology*, *28*, 600–625.
- Tsai, J. L., Levenson, R. W., & Carstensen, L. L. (2000a). Autonomic, expressive, and subjective responses to emotional films in older and younger Chinese American and European American adults. *Psychology & Aging*, *15*, 684–693.
- Tsai, J. L., Ying, Y.-W., & Lee, P. A. (2000b). The meaning of ‘being Chinese’ and ‘being American’: Variation among Chinese-American young adults. *Journal of Cross-Cultural Psychology*, *31*, 302–332.
- Undeutsch, U. (1989). The development of statement reality analysis. In J. C. Yuille (Ed.), *Credibility assessment* (pp. 101–119). New York: Kluwer Academic/Plenum Publishers.
- Vrij, A. (2007). Criteria based content analysis: A qualitative review of the first 37 studies. *Psychology, Public Policy, and Law*, *11*, 3–41.
- Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities*. Chichester, England: Wiley.
- Vrij, A., Akehurst, L., Soukara, S., & Bull, R. (2004). Let me inform you how to tell a convincing story: CBCA and reality monitoring scores as a function of age, coaching, and deception. *Canadian Journal of Behavioural Science*, *36*, 113–126.
- 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 and Human Behavior*, *31*, 499–518.
- Weintraub, W. (1989). *Verbal behavior in everyday life*. New York: Springer.