Running head:

META-ANALYSIS OF COMPUTER-ASSESSED LINGUISTIC CUES TO DECEPTION

Are Computers Effective Lie Detectors?
A Meta-Analysis of Linguistic Cues to Deception

Valerie Hauch[1], Iris Blandón-Gitlin[2], Jaume Masip[3],

and Siegfried L. Sporer[1]

Author Note

[1] Justus Liebig University Giessen, Germany

[2] California State University, Fullerton, USA

[3] University of Salamanca, Spain

Acknowledgments. The authors wish to thank all the authors of primary studies who took the effort to respond to our inquiries.
During the final part of this research, the third author was supported by a grant from the Spanish Ministry of Education, Culture, and Sports, “Programa Nacional de Movilidad de Recursos Humanos del Plan Nacional de I-D+i 2008-2011 (prorrogado por Acuerdo de Consejo de Ministros de 7 de octubre de 2011)” (Ref., PRX12/00209).

Correspondence concerning this article should be addressed to Valerie Hauch or Siegfried L. Sporer, Department of Psychology and Sports Science, Justus Liebig University Giessen, Otto-Behaghel-Strasse 10F, 35394 Giessen, Germany. E-mail: Valerie.Hauch@psychol.uni-giessen.de; sporer@psychol.uni-giessen.de.

Personality and Social Psychology Review

Are Computers Effective Lie Detectors? A Meta-Analysis of Linguistic Cues to Deception

Journal:	*Personality and Social Psychology Review*
Manuscript ID:	PSPR-01-14-004.R2
Keywords:	Applied Social Psychology, Communication, Detection of Deception, Linguistic Cues, Meta-Analysis, Computer Programs

Abstract

This meta-analysis investigates linguistic cues to deception and whether these cues can be detected with computer programs. We integrated operational definitions for 79 cues from 44 studies where software had been used to identify linguistic deception cues. These cues were allocated to six research questions. As expected, the meta-analyses demonstrated that, relative to truth-tellers, liars experienced greater cognitive load, expressed more negative emotions, distanced themselves more from events, expressed fewer sensory-perceptual words, and referred less often to cognitive processes. However, liars were not more uncertain than truth-tellers. These effects were moderated by event type, involvement, emotional valence, intensity of interaction, motivation, and other moderators. Although the overall effect size was small theory-driven predictions for certain cues received support. These findings not only further our knowledge about the usefulness of linguistic cues to detect deception with computers in applied settings but also elucidate the relationship between language and deception.

Keywords: detection of deception, linguistic cues, computer program, meta-analysis

Are Computers Effective Lie Detectors?

A Meta-Analysis of Linguistic Cues to Deception

Deception is an ubiquitous phenomenon, and people at all times have sought to find ways to detect it. Humans have searched for indicators of deception in physiological, nonverbal and paraverbal behavior, and the very content of what people are saying. Since the beginning of experimental psychology, researchers have systematically investigated different types of cues assumed to reveal deception (Benussi, 1914; Freud, 1905; Wertheimer & Klein, 1904; see Bunn, 2012; Grubin & Madsen, 2005; Sporer, 2008, for historical reviews).Despite these efforts, meta-analyses indicate that humans are not very good at discriminating between truths and lies (Bond & DePaulo, 2006). Reasons may lie in the complexity and difficulty of the task, incorrect beliefs about cues and the use of invalid cues, as well as the pervasive biases in decision making (Global Detection Research Team, 2006; Reinhard, Sporer, Scharmach, & Marksteiner, 2011; Vrij, 2008b).

In this meta-analysis, we focus on the use of computers to overcome these limitations. However, we unpretentiously believe the present contribution goes far beyond this goal. Based on a series of theoretical frameworks rooted in cognitive and social psychology, we posed (and tested) specific directional hypotheses concerning the potential utility to detect deception with a number of linguistic cues. Our findings are relevant not only in terms of the potential practical utility of computers to detect deception, but also in terms of basic knowledge about the language of deception and the underlying theories predicting specific linguistic differences between truths and lies.

Human Judgmental Biases

Humans are biased lie detectors. Biases include a reliance on cognitive heuristics (Levine & McCornack, 2001), overestimation of dispositional factors (O'Sullivan, 2003), and an exaggerated focus on nonverbal relative to verbal content cues (Reinhard et al., 2011; Vrij, 2008b). Other researchers have shown that humans are prone to truth or lie biases (Levine, Park, & McCornack, 1999; Meissner & Kassin, 2002; Zuckerman, Koestner, Collela, & Alton, 1984), which are the tendency to judge statements as truthful--or as deceptive--regardless of their actual veracity. It has also been shown that observers' veracity judgments are affected by factors unrelated to the veracity of particular statements, such as the sender's facial appearance (Masip, Garrido, & Herrero, 2003). Likewise, Bond and DePaulo (2006) argue that people hold the stereotype that liars are "tormented, anxious, and conscience stricken" (p. 216), and that they may draw on this stereotype when judging the veracity of other people.

As a possible remedy to overcome these deficiencies in human judgments, physiological psychologists and brain researchers have utilized "machines" like the polygraph, voice stress analyzer, pupillometry, electromyogram, and brain imagery (e.g., EEG, fMRI) to detect deception. In the last 40 years, but particularly most recently, scientists from various fields have also sought to detect deception by analyzing speech content with computers, looking for specific word cues or sentence structures to reveal deception.

A computer system would arguably be less prone to the influence of biases and stereotypes than human judges. There would be virtually no top-down processing. Additionally, online assessment of various deception cues from ongoing interactions or videos can tax the cognitive capacity of human judges and lead to errors. Computers can quickly analyze large amounts of information and provide more reliable data. These are the principal reasons for the appeal of the automatization of lie detection. However, we must not forget that computers do not make choices about definitions of word categories nor about the specific words to be contained in broader categories. Most importantly, computers do not make choices about the *direction* of any particular cue as a lie or truth indicator. It is important to stress that, for a computer to be able to detect deception, the linguistic characteristics to be analyzed must be revealing of deception. Here, in examining what linguistic cues identified with computers differ between truths and lies, we also contribute to our basic understanding about linguistic markers of deception.

Can Computers be Useful to Detect Deception?

In an attempt to identify and quantify linguistic cues to deception, researchers had an (unrealistic) dream: Enter peoples' words into a computer to find out if they are telling the truth or not. In an early study, Knapp, Hart, and Dennis (1974) assessed several linguistic cues using a program called TEXAN on a CDC 6500 mainframe computer. The program analyzed word frequencies without taking contextual meaning into account. Most of the investigated cues significantly differed in the expected direction between truths and lies.

Many years passed until similar but more modern word frequency count approaches were used regularly to deception detection (at least in research contexts). The most common program, called Linguistic Inquiry and Word Count (LIWC; Pennebaker, Francis, & Booth, 2001), was developed to count words in psychology-relevant dimensions across multiple text files. LIWC has been used in numerous domains like personality, health, or psychological adjustment (see Tausczik & Pennebaker, 2010, for a review). LIWC analyzes typed or transcribed accounts on a word-by-word basis, where each word is compared against a dictionary of 2000 pre-selected words allocated to 72 linguistic categories. Although LIWC was not specifically designed to assess deception, Newman, Pennebaker, Berry, and Richards (2003) used it to calculate the percentages of specific linguistic cues in true versus deceptive statements, yielding above-chance accuracy of classifications for different types of lies. Subsequently, researchers from a variety of fields have also applied LIWC with the same purpose (see Appendix C).

Other researchers realized that the methods used ought to be more complex. As a result, specialized programs and algorithms have been developed which are oriented more directly to detecting deception. For example, Agent99Analyzer was created to specifically detect (linguistic cues to) deception in texts and videos (Fuller, Biros, Burgoon, Adkins, & Twitchell, 2006). One of its sub-tools is a natural language processing unit called "GATE" (General Architecture for Text Engineering; Cunningham, 2002; Qin, Burgoon, Blair, & Nunamaker, 2005). Other related automated text-based tools used were "iSkim" or "CueCal" (Zhou, Booker, & Zhang, 2002; Zhou, Burgoon, Nunamaker, & Twitchell, 2004). More specifically, smaller text units are analyzed and integrated in the context of the whole text through examining different levels of human language (e.g., sub-sentential, sentential and discourse processing; see also Zhou et al., 2004). Recently, a growing body of research using machine-learning approaches of natural language processing emerged to detect linguistic cues to deception (Nunamaker, Burgoon, Twyman, Proudfoot, Schuetzler, & Giboney, 2012).

A highly sophisticated program of this kind called "Coh-Metrix" (Graesser, McNamara, Louwerse, & Cai, 2004; McNamara & Graesser, 2012), goes beyond word frequency analysis. Specifically, in analyzing "cohesion relations", Coh-Metrix takes into account meaning and context in which words or phrases occur in texts (http://cohmetrix.memphis.edu). Although not specifically developed to detect deception, Coh-Metrix was recently applied for this purpose (e.g., Bedwell, Gallagher, Whitten, & Fiore, 2011). A somewhat different detection deception software called Automated Deception Analysis Machine ("ADAM"; Derrick, Meservy, Burgoon, & Nunamaker, 2012) focuses on editing processes while typing messages (e.g., backspace, delete, or spacebar) and measures response latencies. The program includes an automated interviewer asking questions from an internal script.

Taken together, various computer programs from different research areas and labs originated in the last 15 years that were either applied to detecting deception or specifically developed for this purpose. The effectiveness of such programs can be better determined with a comprehensive and integrative quantitative analysis of the results on various linguistic cues to deception. This is the focus of the current meta-analysis.

The Importance of Theory

Is this dream of automated lie detection realistic? A quick preview of our results hints to the fragmented nature of the findings from computer studies. Effect sizes in our meta-analysis were coded in a way that *positive* g_us are indicative of *truth*, while *negative* g_us are indicative of *deception*. For 1,093 effect sizes we calculated for 79 linguistic cues, we obtained an approximately normal distribution centering on a mean effect size of g_u = -0.01 (*SD* = 0.37), and a *Mdn* of 0.02. The effect sizes ranged from -1.95 to 1.43 and the first and third quartiles were -0.17, and 0.20, respectively. To get a more accurate picture of the diagnostic usefulness of linguistic markers of deception, we calculated the *absolute* magnitude of all effect sizes, assuming that all were in the expected direction as predicted by a-priori specified hypotheses (Figure 1). The average absolute effect size was 0.26 (*SD* = 0.26) with a *Mdn* of 0.19 (first quartile = 0.09, third quartile = 0.34). This average effect size denotes the maximum possible mean of all cues *if* the results had actually been in the direction predicted. This mean effect size implies that across all studies and cues only small effect sizes were obtained. This suggests that without a-priori theoretical predictions, computer analyses of linguistic cues to deception are a futile exercise. Can larger effect sizes be observed if we classify cues into theoretically meaningful categories and consider possible moderators?

Theoretical Approaches Used to Predict Linguistic Cues to Deception

We cannot provide an exhaustive review of all approaches taken by different research groups. Some authors may prefer to emphasize the role of emotion, arousal and motivation, while communication researchers may look at deception as strategic behavior. We will address some of these alternative interpretations where appropriate. Instead, we focus more on a cognitive and memory-oriented approach, supplemented by social psychological considerations and self-presentation, which help us to pin down the differences in processes involved in telling true stories vs. lies. Hence, we focus on four viewpoints resulting in six research questions: (1) Recalling an experience from episodic memory vs. constructing a lie from semantic memory. Constructing a lie may be more cognitively taxing (Research Question 1) and reduces the certainty with which lies are delivered (Research Question 2). (2) Again drawing on the literature on memory, we discuss the role of emotion and affect in recall of true experiences vs. reporting lies (Research Question 3). (3) We discuss the role of the self as an organizational principle as well as self-presentational strategies and the role of immediacy in communication (Research Question 4). (4) We draw on the reality monitoring framework to derive predictions about sensory and perceptual cues (Research Question 5) and cognitive operations (Research Question 6).

For each question we noted those linguistic cues that would elucidate differences between accounts of truth-tellers and liars, clearly specifying the direction of effect for each cue. Some of the theoretical approaches we discuss elaborate retrieval and construction processes truth-tellers engage in when reporting an event while others focus on lie construction. Furthermore, we developed clear operational definitions for each cue in order to provide consistency in the names and definitions used in different research areas (see Appendices A and B). Most cues investigated could be allocated to one of the six research questions. However, because some cues did not clearly fit in any theory or research question, they were relegated to the miscellaneous question category. Following are the principal research questions.

Research Question 1: Do Liars Experience Greater Cognitive Load?

Telling a lie can be more cognitively demanding than truth-telling, because it involves the execution of a number of concurrent tasks requiring a great deal of mental resources. In general, both liars and truth-tellers must tell a plausible and coherent story that does not contradict their own former statements or facts the observer/interviewer may know about. Also, in some cases lying requires suppressing thoughts about the truth (Gombos, 2006); this may inadvertently preoccupy the speaker's thinking (Pennebaker & Chew, 1985; see also Lane & Wegner's, 1995, model of secrecy). Further, as communication researchers have emphasized, storytellers must monitor their own behaviors and observers' reactions (Buller & Burgoon, 1996). Truth-tellers may also engage in some of these cognitive processes but for liars this task is more difficult because they cannot easily draw on episodic memories. Instead, they must rely on the semantic memory system or on rather nonspecific scripts or schemata (Schank & Abelson, 1977; Sporer & Küpper, 1995).

When constructing a lie, a convincing scenario has to be communicated. However, due to the demands for cognitive resources, a lie may not include the complexities and richness of information that characterize reports of real experiences. In contrast, telling a story about a true event relies on retrieval of experienced events. Although this typically involves reconstruction, and may at times even take increased effort, recall of episodic memories and supporting details is generally rather automatic.

Much research on the cognitive load approach has not been grounded on well-articulated cognitive models of deception (Blandón-Gitlin, Fenn, Masip, & Yoo, 2014). Yet, a few such models have been proposed to specify cognitive processes involved in lie production (for reviews, see Gombos, 2006, and Walczyk, Igou, Dixon, & Tcholakian, 2013). Some of these models (Sporer & Schwandt, 2006, 2007; Walczyk, Schwartz, Clifton, Adams, Wei, & Zha, 2005; Walczyk, Harris, Duck, & Mulay, 2014; Walczyk et al., 2013) have invoked Baddeley's (2000, 2006) working memory model, which involves transferring information from long-term memory to an episodic buffer in working memory. While this should facilitate truth-telling, it should also make lying more difficult (see, e.g., Walczyk et al., 2005, 2013, 2014).

Does research support the cognitive load assumptions? Numerous recent studies (for review, see Vrij & Granhag, 2012) have provided indirect evidence by experimentally increasing a storyteller's task demands. This has elicited more discernable cues to deception than in control, lower cognitive load conditions. Note, however, that manipulating "cognitive load" is not equivalent to assessing the cognitive mechanisms postulated as a function of such manipulations (Blandón-Gitlin et al., 2014). More direct (and revealing) evidence comes from behavioral studies using response latencies and other indices of cognitive load (e.g., Debey, Verschuere, & Crombez, 2012; Johnson, Barnhardt, & Xhu, 2004; Walczyk et al., 2005; for a summary, see Walczyk et al., 2013). There is even evidence from brain imaging studies (e.g., Abe, 2009; Christ, Van Essen, Watson, Brubaker, & McDermott, 2009) showing that telling lies, particularly those involving short responses, requires greater involvement of and access to key mental resources than truth-telling (Gamer, Bauermann, Stoeter, & Vosse, 2008).

Cues to deception theoretically connected to the cognitive load perspective have been found in previous meta-analyses, particularly for nonverbal and paraverbal behaviors (DePaulo, Lindsay, Malone, Muhlenbruck, Charlton, & Cooper, 2003; Sporer & Schwandt, 2006, 2007). In comparison with truth-tellers, liars had longer response latencies, tended to communicate shorter stories, made more speech errors, nodded less, and displayed fewer hand, foot, and leg movements. Particularly relevant for the analysis of linguistic markers are findings on verbal content cues that demonstrate that compared to true accounts, deceptive accounts appear less plausible, coherent and detailed while including more phrase and word repetitions. These indices can be signs of the experience of cognitive load either from a taxed system (e.g., longer response latencies) or because of liars' strategies to reduce cognitive load (Walczyk, Mahoney, Doverspike, & Griffith-Ross, 2009).

Predictions. From a cognitive load/working memory perspective, we predict that compared to true accounts, false accounts will be (a) shorter as indicted by *word* and *sentence quantity* cues, (b) less precisely elaborated as indicated by fewer *content words* (expressing lexical meaning), a lower *type-token ratio* (number of distinct content words, e.g., *house, walk, mother*) divided by total number of words), and *shorter words* (i.e., less than six letters; average word length), (c) involve less complex stories as indicated by fewer *verbs,* fewer *causation* words (*because*, *effect*, *hence*) and fewer *exclusive words* (*but*, *except*, *without*), and (d) include more *writing errors* (possibly moderated by mode of production [orally telling a lie, hand writing, or typing]). (For a list of the operational definition of all cues included see Appendices A and B.)

From a different perspective, based on DePaulo's self-presentational perspective (DePaulo et al., 2003), one would expect that liars are less likely than truth-tellers to take their credibility for granted and therefore may take a greater effort and deliberately edit their communication (cf. Derrick et al., 2012). Note, however, that this editing process will also usurp cognitive resources detracting from successful lie constructions.

Research Question 2: Are Liars Less Certain Than Truth-Tellers?

DePaulo et al. (2003) contend that deceptive self-presentations are not as convincingly embraced as truthful ones. This may be a result either of the speakers' moral scruples, which may lead them to feel guilty or ashamed when lying, or of liars not having as much personal investment in their claims as truth-tellers. The psychological closeness or distance between a speaker and his or her message might be reflected in language (Wiener & Mehrabian, 1968). Liars should display more linguistic markers indicative of psychological detachment than truth-tellers (Buller, Burgoon, Busling, & Roiger, 1996; Kuiken, 1981;

Wagner and Pease, 1976; Zhou, Burgoon, Nunamaker, & Twitchell, 2004; Zhou, Burgoon, Twitchell, Qin, & Nunamaker, 2004). Indeed, in their meta-analysis DePaulo et al. (2003) found that liars were verbally and vocally less involved and more verbally and vocally uncertain than truth-tellers but observed no reliable differences for tentative constructs and shrugs. Uncertainty words have been proposed as markers of psychological distance between a speaker and his or her account (e.g., Kuiken, 1981). Thus, liars' accounts should contain more uncertainty words than truth-tellers' accounts.

It may also be the case that deceivers withhold information not to give their lies away. Indeed, research shows that when lying to conceal their transgressions, people indicate that they try not to provide incriminating details (Hartwig, Granhag, & Strömwall, 2007; Masip & Herrero, 2013), and try to keep the story simple (Strömwall, Hartwig, & Granhag, 2006) or vague (Vrij, Mann, Leal, & Granhag, 2010). DePaulo et al. (2003) found liars to be significantly more discrepant/ambivalent than truth-tellers. Therefore, liars might provide vague, ambiguous, or uncertain replies in order not to expose their lies (Buller et al., 1996; Cody, Marston, & Foster, 1984).

Predictions. From these perspectives, it is expected that liars will be less certain and definite than truth-tellers. Consequently, deceptive accounts should contain fewer *certainty words* (*always*, *clear*, *never*) and more *tentative words* (*guess*, *maybe*, *perhaps*, *seem*) and *modal verbs* (*can*, *shall*, *should*) than truthful accounts. (It should be noted that *modal verbs* also include the verb "must" that expresses more certainty and purposiveness whereas all other modal verbs indicate more uncertainty).

It may be argued that liars are aware that uncertainty indicates deception and thus may strategically incorporate certainty indicators to evade detection (e.g., Bender, 1987). However, research does not support this contention. To our knowledge, around ten reports have been published so far on liars' and truth-tellers' strategies to be convincing (for a brief review, see Masip & Herrero, 2013). Only rarely has certainty (or any related construct) emerged as a strategy, and in these instances it has been mentioned (a) only infrequently, and (b) equally often by liars and truth-tellers (e.g., Hines, Colwell, Hiscock-Anisman, Garrett, Ansarra, & Montalvo, 2010: "admit uncertainty"; for an exception see Strömwall et al., 2006).

Research Question 3a: Do Liars Use More Negations and Negative Emotion Words?

Emotional approach.[1] When people lie, they may experience feelings of guilt and fear of getting caught (Ekman, 1988, 2001).[2] Even when telling everyday lies of little consequence, people report feeling uncomfortable (DePaulo et al., 2003). Vrij (2008a) also noted that liars might make negative comments or use negative words that reflect negative affect induced by guilt and fear.

Numerous studies have shown that arousal is associated with specific emotions (see the meta-analysis by Lench, Flores, & Bench, 2011), some of which are likely to be experienced by liars, such as guilt and fear of punishment (Ekman, 2001; Zuckerman, DePaulo, & Rosenthal, 1981). These emotional states may elicit specific nonverbal and verbal cues to deception (see DePaulo et al., 2003; Sporer & Schwandt, 2006; Vrij, 2008a). Recent studies have used brain-imaging technology to specifically investigate the role of emotion in deception (for a review see Abe, 2011). For example, Abe, Suzuki, Mori, Itoh, and Fujii (2007) found that neural structures associated with heightened emotions were also uniquely associated with deceiving an interrogator, and that self-reported feelings of immorality (sense of sin) and anxiety were higher in deceptive conditions than in truth-telling conditions. These results support the notion that deception is associated with negative emotions.

Predictions. From an emotional approach perspective, we predict that compared to true accounts, lies will include (a) more *negation* words (*no*, *never*, *not*) because these reveal a more defensive tone or denial of wrongdoing, which is likely to be accompanied by negative emotions of the liar, and (b) more words denoting overall *negative emotions (enemy, worthless, skeptic), anger (hate, kill, weapon), anxiety (unsure, vulnerable) and sadness (tears, useless, unhappy)*.

Research Question 3b: Do Liars Use Fewer Positive Emotion Words?

Research on autobiographical memory suggests that people's emotional appraisal of past events tends to be positively biased (Walker, Vogl, & Thompson, 1997). One mechanism by which this bias occurs is a tendency for emotions associated with negative-event memories to fade faster than emotions associated with positive-event memories (Walker, Skowronski, & Thompson, 2003). In a review of this research, Walker and Skowronski (2009) suggest that this *fading affect bias* leads people to generally remember events less negatively regardless of the original affect associated with the event. This effect is not due to forgetting of event details, as the accuracy of the memories is comparable for negative and positive events. It is the memory of the emotional intensity associated with the event that fades, with negative events fading at a faster rate than positive events.

Predictions. Because truth-tellers have a specific memory of the event, whereas liars cannot draw on such an episodic memory, we predict that compared to true accounts, lies will contain fewer words denoting *positive emotions* (happy, pretty, good) or *feelings (luck, joy)*.

Research Question 3c: Do Liars Express More or Less Unspecified Emotion Words?

Many researchers from different fields, such as social psychology, psychology and law, or computer linguistics (e.g., Ali & Levine 2008; Fuller et al., 2006, Newman et al., 2003), have investigated the frequency of occurrence of emotional and affective terms in true and deceptive accounts without taking the valence of these emotions into account. Therefore, we decided to also investigate the cues of *unspecified emotions* (positive and negative) and *pleasantness* or *unpleasantness* of the story despite the lack of theoretical specification of the direction in the original studies. Predictions could be derived from a social psychological perspective. Depending on the seriousness of a lie, from a trivial lie in everyday life to high stake lies, the situation may become increasingly emotional. Hence, one would predict higher frequencies of *unspecified emotion* words in lies than in truths.

Research Question 4: Do Liars Distance Themselves More From Events?

In the preceding section, we have assumed that people are more likely to experience different types of negative emotions when telling a lie. Given such negative experiences and emotions, from DePaulo et al.'s (2003) self-presentational perspective we further assume that liars will distance themselves more from the story being told, and, relatedly, will be less forthcoming than truth-tellers (see also Research Question 2 on certainty cues above). Possible linguistic indicators for this assumption are personal pronouns, cues to responsibility and verb tense shifts. To clarify the predictions of specific cues we present them within the theoretical accounts of immediacy, self-organization, egocentric bias, and narrative conventions.

Immediacy. A possible way to express ownership and take responsibility for an action or event is to tell a story from a first-person perspective, where the sender is reporting an event where he/she is the actor, not an observer-bystander. Evidence for this assumption comes from the long tradition of research on verbal and nonverbal communication which has investigated *immediacy* as a cue to truthful messages (Cody et al., 1984; Knapp et al., 1974; Kuiken, 1981; Mehrabian, 1972; Wagner & Pease, 1976; Wiener & Mehrabian, 1968; Zhou, Burgoon, Nunamaker et al., 2004; Zhou, Burgoon, Twitchell et al., 2004). In these studies, one aspect of immediacy has been operationalized as the psychological distance between the speaker and his/her communication. More specifically, immediacy can indicate the degree to which there is directness and intensity between the communicator and the event being communicated (Wiener & Mehrabian, 1968, p. 4). Taking this aspect of the definition of immediacy, deception researchers consider nonimmediacy as an indicator of deceptive communication by way of the speaker distancing from his/her own statement (e.g., Buller et al., 1996; Kuiken, 1981; Wagner & Pease, 1976; Zhou et al., 2004).

However, evidence for nonverbal and verbal indicators of the relationship between immediacy and deception is mixed. In the meta-analysis by DePaulo et al. (2003) there were no significant effects for self- or other-references, but more general indices of *verbal* immediacy (all categories) as well as verbal and vocal immediacy (impressions) were observed significantly more frequently or to a higher extent in truthful than fabricated messages. This latter effect appeared to be stronger when immediacy was measured subjectively than when assessed via more objective measures.

The self as an organizational structure. Another line of research we consider is social psychological theorizing on social memory, which has emphasized the role of the self as an organizational structure. In fact, one of the primary distinctions between episodic and autobiographical memory is that the self provides an organizing principle, which relates experiences to one's self-schema. Experimental evidence comes from research on the self-reference effect (Rogers, Kuiper, & Kirker, 1977), which demonstrated that information is particularly well remembered when it has been encoded in relation to oneself, or when the person plays an active, rather than passive role (e.g., Slamecka & Graf, 1978). Variations on this theme are discussed under ego-defensive, self-serving, egocentric or egotistic biases (see Greenwald, 1980). Greenwald (1980) has gone as far as referring to the self as a "totalitarian ego" that puts itself in the foreground, assuming a central role and ownership when talking about self-experienced past events and actions. This prevailing tendency should lead to more frequent uses of first-person pronouns (*I*, *me*, *we*, *us*, *our*, etc.) when telling the truth relative to lying.

However, while the *egocentric bias* may play a role when reporting (complex) autobiographical events, it may be restricted to positive outcomes, and reversed for negative outcomes (Greenwald, 1980). Also, the so-called "better than average effect" refers to the tendency to evaluate oneself more favorably than an average peer (e.g., Brown, 2012). For instance, 70% of high school seniors estimated that they had above average leadership skills, whereas only 2% said their leadership skills were below average (College Board, 1976–1977). Another example of the positive outcome bias is a classic study by Bahrick, Hall, and Berger (1996; see also Bahrick, 1996) who found that students accurately recalled better high school grades than worse ones. Relatedly, in a classical study on the self-enhancing bias by Cialdini, Borden, Thorne, Walker, Freeman, and Sloan (1976, Experiment 2) college students not only donned their school colors on Monday after their team had won, but also identified, or distanced, themselves by use of different personal pronouns ("we won"; "they lost"). This suggests that first-person pronouns and statements of personal responsibility will be more prevalent among truth-tellers than liars, particularly for positive outcomes.

Predictions. In summary, from different theoretical perspectives we assume more frequent use of *first-person pronouns*, and less frequent use of *third-person pronouns* for reports of self-experienced events. Self-experienced events should also be characterized by more statements of own responsibility, at least for positive outcomes. This prediction is more likely to hold for *first-person singular* than *first-person plural* because the plural may designate both the group the storyteller belongs to, and identifies with, as well as a communication partner who acts as an antagonist in an interaction (e.g., "we quarreled"). Thus, with plural pronouns, ownership and responsibility are less clear-cut than with singular pronouns. On the other hand, *passive voice* or *generalizing terms* in phrases like "one has to..." or "everybody does this..." signal less personal involvement and hence should be found more frequently in lies than truthful accounts.

Narrative conventions and verb tense shifts. Communication about past events follow narrative conventions (acquired during childhood) that require the storyteller to talk about who, what, when, where, and why (Brown & Kulik, 1977; Neisser, 1982) and to adhere to a temporal structure (Bruner, 1990). Anecdotal evidence from research on autobiographical memory for significant life events shows that people sometimes switch from telling a story in the past tense to the present tense at crucial moments of the event (Pillemer, Desrochers, & Ebanks, 1998). In many of these examples, it appears that the protagonist is reliving the past event, describing his or her sensory and perceptual experiences, making the accounts to appear more vivid (cf. the reality monitoring approach described in Research Question 5). Although present tense may be less concrete than past tense when it refers to repeated or routine actions (e.g., "I [usually] go to church on Sunday" versus "I went to church on Sunday"), when talking about a specific past event present tense is more vivid than past tense. Whether verb tense shifts occur involuntarily or unconsciously, or are strategically used by skillful storytellers (like fiction writers) to communicate intensity and feeling to a recipient, cannot be answered by these archival type studies, nor by our meta-analyses.

Predictions. We expect reports of true events to be more likely to contain *present tense verbs* than lies, at least in accounts of personally significant events. For other types of lies, this prediction may not hold. The live character of these narratives may also diminish with repeated retellings of a story. Conversely, lies should contain more *past tense verbs* than true accounts.

Research Question 5: Do Liars Use Fewer (Sensory and Contextual) Details?

Reality monitoring framework applied to deception. The reality monitoring model by Johnson and Raye (1981) describes how individuals differentiate between externally generated memories of actual experiences versus memories of internally generated events that involve thoughts, fantasies, or dreams. In contrast to imagined events, experienced events are encoded and embedded in memory within an elaborate network of information that typically includes more perceptual details, contextual and semantic information. Conversely, internally generated memories are characterized by cognitive inferences or reasoning processes.

People differentiate between their own external and internal memories on the basis of these phenomenal characteristics (Johnson, Hashtroudi, & Lindsay, 1993), and similar features are also useful to differentiate between accounts of external and internal memories of *other people* (an attribution process that has been tagged "interpersonal reality monitoring"; Johnson, Bush, & Mitchell, 1998; Johnson & Suengas, 1989; Sporer, 2004; Sporer & Sharman, 2006).

Deceptive accounts can be characterized as representing internally generated memories, because in a deceptive situation people imagine the event at the time of its construction (Sporer, 2004). Even if people lie by borrowing from actual experience, the time and place or the context in which the event occurred may be changed during construction (Sporer, 2004; Vrij, 2008a). Therefore, even partially true deceptive accounts may lack the typical characteristics of true accounts. With these considerations in mind, researchers have extrapolated from the reality monitoring model to make predictions about specific sets of criteria that may discriminate between true and deceptive accounts (e.g., Granhag, Strömwall, & Olsson, 2001; Sporer, 1997; for reviews see Masip, Sporer, Garrido, & Herrero, 2005; Sporer, 2004; Vrij, 2008a). DePaulo et al.'s (2003) meta-analysis, which only included a few studies available then, showed small and nonsignificant effects sizes for reality monitoring criteria. However, in a more comprehensive review of studies, Masip et al. (2005) found that some of the reality monitoring criteria involving perceptual processes, contextual (including time) information, and realism/plausibility of the story were useful to discriminate between truth and deception.

Predictions. From a reality monitoring perspective, we predict that compared to true accounts, false accounts will (a) contain fewer perceptual details as indicated by *sensory and perceptual word* cues (*taste, touch, smell*), (b) be less contextually embedded as indicated by *space* (*around, under*) *and time word* cues (*hour, year*), and (c) include fewer descriptive words as indicated by *prepositions* (*on, to*), *numbers* (*first, three*), *quantifiers* (*all, bit, few*), *modifiers* (*adverbs and adjectives*), *and motion verbs* (*walk, run, go*). This latter set of cues involves words that describe events and actions in the story in more specific terms (e.g., "I took every short cut to get to work"). The lack of these words (e.g., "I went to work") would make the account seem less real or vivid as would be predicted from the reality monitoring perspective (Sporer, 1997, 2004).

Research Question 6: Do Liars Refer Less (yes, Less!) Often to Cognitive Processes?

The reality monitoring approach, unlike other verbal-content cues based credibility assessment procedures, such as Criteria-Based Content Analysis (CBCA, Steller & Köhnken, 1989), does not only contain "truth criteria" (e.g., spatial and time details), but also one lie criterion. Specifically, reality monitoring predicts that references to internal processes at the time of the event (cognitive operations like reasoning processes) should be more likely contained in imagined than in self-experienced events. Applied to detecting deception, researchers have consequently postulated that references to cognitive operations can be used as a lie criterion (Sporer, 1997; Vrij, 2008a).

However, empirical evidence regarding this proposition is mixed. Perhaps, depending on the operationalization of this construct, some studies have found more references to cognitive operations in *lies* (e.g., Vrij, Akehurst, Soukara, & Bull, 2004), many studies have found *no differences* (e.g., Sporer & Sharman, 2006; 14 out of 19 studies reviewed in Vrij, 2008a), and some studies have found reliably more references to internal processes (like memory processes and rehearsal as well as thoughts) in *true* accounts (Granhag, Strömwall,

& Olsson, 2001; Sporer, 1998; Sporer & Walther, 2006; Vrij, Edward, Roberts, & Bull, 2000).

From a different perspective, some thirty years of research on autobiographical memory has emphasized the associative nature of memories. Recollecting (personally significant) life events involves not only the conscious utilization of retrieval cues but also cross-referencing to supporting memories related to the event in question. It also involves rehearsal processes, which are important determinants of remembering (Conway, 1990). These processes can also be subsumed under cognitive operations. To the extent that studies on deception involve complex (autobiographical) events, like being questioned about a crime or reporting an alibi, such retrieval processes and supporting memories (cf. the Criteria-Based Content Analysis criterion "External Associations") are likely to be used and mentioned when recalling true events (e.g. "I know it was the day before Easter because Good Friday was my birthday.").

Finally, there is empirical evidence from several studies that cognitive operations are *positively* correlated not only with other reality monitoring criteria (Sporer, 1997, 2013) but also with many Criteria-Based Content Analysis criteria like "External Associations", "Own Psychological Processes", "Spontaneous Corrections" or "Doubts about one's own Testimony", loading on a common underlying factor (Sporer, 2004, Table 4.4). All of these criteria are assumed to indicate truthfulness.

Predictions. Consequently, we predict that linguistic cues referring to cognitive operations including memory processes are more likely to be found in truths than in lies. The two cues under this research question are *cognitive processes* (*cause, ought*), and *insight words* (*think, know, consider*).

Miscellaneous Category

Because many linguistic cues were investigated without a specific theoretical background or directed predictions, we created a miscellaneous category including linguistic cues analyzed in more than five studies (e.g., *inhibition*, *social processes*, *health, sports*; see Appendix B).

Hypotheses for Moderator Variables

It would be unwise to assume that the above predictions will hold across all types of lies, motivation, level of interaction, production mode, and other contextual factors. Hence, we conducted a series of moderator analyses within the theoretical frameworks provided above.

Event type and personal involvement. Across studies, senders described events or attitudes that differed in terms of personal involvement. We organized the studies into three categories. In the "Attitude/liking" paradigm, senders described their attitude towards a specific topic or person they like or dislike. In the "First-person experience" paradigm senders experienced a staged event or mock crime, described a personal life event, or were involved in a real criminal case. Lastly, the "Miscellaneous" category included studies where participants solved a problem, performed a specific task, or described a video scene.[3] We do acknowledge, however, that some attitudes/liking studies may also reflect high involvement but this would work against our hypothesis.

We argue that the higher the personal involvement in the event the higher the cognitive load (for example, due to a preoccupation with an interaction partner's reactions) and arousal (negative or unspecified emotions) will be when telling a lie. Also, liars might express more uncertainty terms or try to distance themselves more from events when their personal involvement is high. In other words, we expect the effects under Research Questions

1, 2, 3a, 3c, and 4 to be larger for the "First-person experience" compared to the "Attitude/liking" or the "Miscellaneous" paradigms in the aforementioned direction.

Emotional valence. The topics or events senders were asked to talk about were classified as positive (e.g., holidays), neutral (e.g., task performance), or negative (e.g., confession of wrongdoing) in nature. If we assume that more negative emotions accompany telling a negative rather than a neutral event, liars should express even more negative emotion words when the event is negative (Research Question 3a). Also, we assume that the amount of unspecified emotion words (Research Question 3a) will be higher when the event is *not* neutral. Moreover, cognitive load might also be higher because senders have to deal with additional negative emotions that may induce concern, leading to a decrease in *word count* and *diverse* and *exclusive words* (Research Question 1--Cognitive Load: cues 01, 02, 03).

Also, if liars are more negatively involved in their story, they could appear more uncertain (Research Question 2--Certainty) and try to distance themselves more using less *self*- and more *other-references* (Research Question 4--Distancing). In summary, we hypothesized that effect sizes under Research Questions 1, 2, 3a, 3c, and 4 would be highest (in the expected direction) if the emotional valence was negative rather than neutral (or positive).

Intensity of interaction. The degree of interactions between the storyteller and another person varies widely in deception detection research (Vrij & Granhag, 2012). We differentiated four interaction levels: (a) no interaction: participants are only given a written or spoken instruction; (b) computer-mediated communication: participants are communicating via connected computers (e.g., only by *typing* words in studies included); (c) interview: interviewees are simply responding to questions from an interviewer (one-way direction); and (d) person to person interactions: sender and receiver are present in person and interacting bidirectionally.[4] We hypothesized that with increasing intensity of interactions from (a) to (d) (cf. Buller & Burgoon, 1996), effects would become stronger under Research Questions 1, 2, 3a, 3c, and 4.

Motivation. Researchers varied the level of motivation for their senders to appear credible. Some researchers did not motivate their senders at all, some others tried to motivate them with incentives or written instructions, and still others used accounts from real criminal cases, where the motivation to appear credible must have been high due to real consequences for getting caught (*high-stake lies*; cf. DePaulo et al., 2003).

DePaulo and Kirkendol (1989, p. 54) postulated the *motivational impairment effect*, according to which highly motivated liars try to control their expressive behaviors to appear credible, but they are only successful in doing so with their *verbal* behavior, while their *nonverbal* behavior appears disrupted. In other words, liars' nonverbal behavior should be impaired whereas their verbal behavior (i.e., the content of messages) should be improved. DePaulo, Lanier, and Davies (1983) provided support for these hypotheses, as highly motivated liars were easier to detect in the visual or audiovisual conditions, but less successfully detected in the verbal (transcript) condition (there was no difference in the audio-only condition).

Assuming that the motivational impairment effect also applies to linguistic cues as a form of verbal behavior, we hypothesized that highly motivated liars might try harder to control their words, so differences between liars and truth-tellers should become smaller under Research Questions 1, 2, 3a, 3c, and 5.

Production mode. Participants' accounts were either handwritten, typed on a keyboard, or spoken (and audio- or videotaped). Horowitz and Newman (1964) proposed that, in general, speaking is easier than writing, because speakers have more liberty and feel less inhibited than writers. Also, writing involves more deliberateness (see also Hancock, Woodworth, & Goorha, 2010) and more serious commitment. Horowitz and Newman found support for their hypothesis in that speaking is more productive and elaborative than writing. This resulted in more words, more phrases and more sentences when speaking than when writing. More recently, Kellogg (2007) hypothesized that writing is slower and less practiced than speaking and thus results in higher demands on working memory. He found that accounts of a recalled story were more complete and more accurate when spoken than written (cf. also Sauerland & Sporer, 2011).

Hence, we hypothesized that liars produce even fewer *words*, *diverse words*, and *sentences* (Research Question 1--Cognitive Load) when writing than speaking due to an increased cognitive load and decreased working memory capacity. Furthermore, liars should also use fewer (*sensory* and *contextual*) *details* when writing than speaking compared to truth-tellers (Research Question 5; see Elntib, Wagstaff, & Wheatcroft, 2014, for a recent empirical investigation of this issue). Regarding emotion-related cues (under Research Questions 3a, and 3c), we hypothesized that liars use more *negative* and *unspecified emotion words* than truth-tellers when speaking than when writing, because emotions might be expressed more directly and frequently in direct speech.

An empirical issue for studies involving writing is whether handwriting or typing comes easier. Therefore, we separated written accounts into hand-written vs. typed for our moderator analysis. Unfortunately, we do not know the level of typing skill of participants.

To sum up, differences between liars and truth-tellers should be more pronounced in written (typed or handwritten) compared to orally given accounts for linguistic cues under Research Questions 1 (cognitive load) and 5 (details), whereas for emotion-related cues (Research Questions 3a, and 3c), the effect sizes should be larger if stories were spoken than written.

Program type. Researchers from various fields used different computer programs to analyze deceptive and truthful accounts. The most common one is *LIWC*. Although it is a general program (i.e., not specifically designed to detect deception), we separated it from other general programs such as *Coh-Metrix* or *WordScan*. This is because *LIWC* was used in a disproportionally large number of studies. Other software, such as *Agent99Analyzer* or *Automated Deception Analysis Machine*, were specifically developed to detect deception. We hypothesized that studies applying deception-specific programs should yield stronger effects for any linguistic cue than studies using *LIWC* or any other general program based on simple word counts.

Publication status. The tendency that studies with nonsignificant findings are less likely to be written, submitted, and accepted for publication in peer-reviewed journals, is referred to as *publication bias* (Cooper, 2010; Sutton, 2009). In short, the publication of a study may partially depend on its results rather than on its theoretical or methodological quality (Rothstein, Sutton, & Borenstein, 2005). One method to statistically quantify a publication bias is to compare the effect sizes of published and unpublished studies (see Appendix E in supplemental online materials); another is to test for the association between effect sizes and sample sizes (Levine, Asada, & Carpenter, 2009).

Experimental design. We also assessed *experimental design* as a moderator (between- vs. within-participants), assuming larger effects for the latter (see results in Appendix F in supplemental online materials).

Goals of the Meta-Analysis

The main goals of our meta-analysis were (a) to provide a comprehensive set of operational definitions for each linguistic marker, (b) to offer an elaborate theoretical background in order to specify directed predictions for each cue, (c) to provide a quantitative and comprehensive synthesis of linguistic cues to deception assessed with computer programs obtained from interdisciplinary research areas, and (d) to analyze the influence of important theoretical and methodological moderator variables on the outcome of linguistic cues to deception.

Method

Inclusion and Exclusion Criteria

Studies had to meet the following eligibility criteria to be included in our meta-analysis: (1) Use of software to locate linguistic cues; (2) Reports of specific linguistic cues (not just paraverbal/paralinguistic or nonverbal or physiological cues); studies that reported word counts only, but no other linguistic cues were excluded[5]; (3) Independence of data sets: when analyses of the same data set of transcripts and cues were reported in multiple publications, we only included the source published in the journal with the highest publication standard [e.g., peer review] and excluded the other source(s) to ensure independence of all data sets; and (4) Sufficiency of data to calculate effect sizes (see *Effect Size Measure* section below). Furthermore, (5) whenever a field study with statements from real criminal cases met the aforementioned criteria (e.g., ten Brinke & Porter, 2012), special care was taken to assure ground truth had not been established solely on the basis of a court verdict, but in addition from more than one type of external and independent source of evidence (e.g., physical evidence, witness statements, confessions, etc.). However, these studies should be treated with caution because linguistic aspects of the account may have affected the final case disposition (e.g., lie or truth).

Literature Search and Study Retrieval

As a first step, we searched through reference lists of most relevant studies or reviews (e.g., DePaulo et al., 2003; Newman et al., 2003; Tausczik & Pennebaker, 2010; Zhou et al., 2004). Next, several exhaustive literature searches were conducted from September 2011 to February 2012 in the most important psychological research literature databases, such as the

Social Sciences Citation Index (with cited reference search), PsycInfo, Dissertation Abstracts, and Google Scholar, examining articles published between 1945 and February 2012.

The combination and permutations of four keyword clusters were used: (a) decept*, deceit, lie; (b) verbal, linguistic, language; (c) automatic, computer, software, artificial. These searches resulted in 948 published and unpublished articles, which were reduced to 394 after removing duplicates. Then, the inclusion and exclusion criteria were carefully applied. This reduced the number of articles to 99, from which we still had to exclude 54 for different reasons (Appendix G in supplemental online materials), mostly incomplete reporting of data necessary for our analysis. This resulted in 44 relevant data sets that met all inclusion criteria.

Linguistic Cues to Deception

A total of 202 linguistic cues were extracted from the articles and sorted based on their name and operational definition (if available). In some cases, we merged cues with different names that had very similar operational definitions. For example, *type-token ratio*, *unique words*, *lexical diversity*, or *different words*, were all similarly operationally defined and refer to the same construct. We chose the name most commonly used (e.g., *type-token ratio* in the prior example).

All linguistic cues had to be calculated as a ratio of all other words (except raw frequencies of words, verbs and sentences), and had to be investigated in at least $k = 4$ hypothesis tests. This resulted in 79 linguistic cues of which 50 were allocated to one of the six research questions based on their content and theoretical meaning. The remaining 29 cues could not be allocated to a theory or one of the research questions, and were assigned to the Miscellaneous category. All linguistic cues, with all of their names and final operational definitions, are listed in Appendices A and B.

Effect Size Measure

As an effect size measure we used Hedges's g_u (Hedges, 1981; Borenstein, 2009; Lipsey & Wilson, 2001), an unbiased estimator of the standardized mean difference (Cohen's *d*). Here, it is the standardized mean difference of the average frequency or ratio for each linguistic cue between deceptive and true accounts. If a specific linguistic cue occurred more often during deception than truth, g_u has a *negative* sign. If it occurred more often during truth than deception, g_u was assigned a *positive* value. To calculate g_u, we coded means, standard deviations, and *n*s separately for deceptive and true stories. If this information was not given, other appropriate measures (*t*- or *F*-values with 1 degree of freedom in the numerator, or *p*-values) were coded (for formula collections see Borenstein, 2009; Lipsey & Wilson, 2001).

If no relevant statistical data were available, we e-mailed the researchers to request them. In some instances, there may be discrepancies between the effect sizes reported here and those in the original articles. Reasons for such differences are that some authors provided us with more (differentiated) data, that we sometimes chose specific subgroups for the analyses, or calculated the average effect size across subgroups, as explained in more detail under *Meta-Analytic Techniques* below.

Independent Variables and Moderator Variables

After coding typical study characteristics (e.g., study ID, author names, year of publication, number of senders and gender, etc.), we coded for information that defined the moderator variables or further independent variables of potential interest. These were: Publication status (e.g., published, thesis, etc.), type of computer program (LIWC; other general programs like Wordscan, Microsoft Word, or Coh-Metrix; or specific programs like ADAM (Automated Deception Analysis Machine), Agent99Analyzer, GATE (General Architecture for Text Engineering), iSkim, CueCal, or Connexor), language of accounts, theory presented (if any), cue selection (a priori, reported all or significant cues only), age of the senders, experimental design (between- or within-participants), preparation time, event type, event valence, interaction between sender and receiver, mode of production, and type/level of motivation to lie successful.

Coding Procedures and Intercoder Reliability

Two trained raters coded all dependent and independent variables from the articles with a standardized coding manual. After discussing two articles as examples and agreeing on order of article review, each coder worked independently. For eleven continuous variables, inter-coder reliabilities were highly satisfactory, with all coefficients ranging from Pearson's r = .86 to r = 1.0 (except for preparation time: r = .77). For eight categorical variables, inter-coder reliabilities were excellent, with Cohen's *kappa* (Cohen, 1960) ranging from .75 to 1.0. For six additional categorical variables, Cohen's *kappa* ranged from .51 to .67, which was still a fair to good agreement (according to Fleiss, 1981). The few disagreements were resolved by discussion between the two coders. Final coding decisions of the moderator variables for each study are displayed in Appendix C.

Meta-Analytic Techniques

Dependencies of effect sizes. In some studies, in addition to accounts' truth status, other independent variables were manipulated as *between*- or *within*-participants factors and the data were reported separately for these subgroups (e.g., Schelleman-Offermans & Merckelbach, 2010: high- vs. low-fantasy-proneness). In studies with additional *within*-participants factors, dependency was avoided by calculating effect sizes separately for each subgroup and averaging them to ensure that only one effect size per study per linguistic cue was included (Lipsey & Wilson, 2001). In two other studies, a second *between*-participants factor (Ali & Levine, 2008: denials or confessions; Qin et al., 2005: text-chat, audio, face-to-face) was examined; here we included each of these subgroups (with *different* stimulus persons) as independent data sets.

Superordinate categories and sub-cues. Sometimes a linguistic category of cues had differentiated effect sizes that seemed to represent a single construct. As an example, we defined cue 19 with the superordinate category ("umbrella term") *positive emotions and feelings* including results from *positive emotions only* and *positive feelings only*. In studies using LIWC 2001, positive feelings and positive emotions/affects are treated as two different linguistic cues--and the data are reported separately for each (in LIWC 2007, they are combined). To ensure that only one effect size per construct is included, we combined sub-cues to a superordinate category (by averaging their effect sizes). However, we also calculated separate meta-analyses for each of these sub-cues (here: cue 19.1 *positive emotions only* and 19.2 *positive feelings only*) to investigate whether the results are more differentiated, or if merging these cues was justifiable. The same procedure was applied to cue 18 *negative emotions* and to cue 28 *sensory-perceptual processes* (see Table 1). These superordinate categories make results from LIWC more comparable with studies using other computer programs that did not differentiate between different sub-cues (e.g., *anger*, *anxiety*, *sadness*).

Weighted average effect size. For each of the 79 linguistic cues, individual meta-analyses under the fixed-effects model (Lipsey & Wilson, 2001; Shadish & Haddock, 2009; Sporer & Cohn, 2011) were calculated. Average effect sizes were weighted by the inverse of the variance to give more weight to studies with larger samples. For six studies the total number of accounts was extremely large. To avoid unjustified extra-ordinary large weights we adjusted the number of total accounts for these studies (see *Results* section).

Homogeneity of effect sizes. We report both the homogeneity test statistic Q (Lipsey & Wilson, 2001) and the descriptive homogeneity statistic I^2 (Higgins & Thompson, 2002;

Shadish & Haddock, 2009). In rare cases where I^2 resulted in a negative value, it was set to 0. In case of heterogeneity, outlier and moderator analyses were conducted.

Outlier analysis. To test for the presence of outliers, we applied the two methods recommended by Hedges and Olkin (1985, Chapter 12, and programmed by the fourth author). The number of outliers did not exceed 15% of the total number of effect sizes to avoid an artificial restriction of the variance between effect sizes. If outliers were detected, we calculated each meta-analysis with and without the outliers as sensitivity analyses (Greenhouse & Iyengar, 2009). Due to space limitations, we only report results without outliers in Table 1 (results with and without outliers are displayed in Appendix H in supplemental online materials).

Moderator analyses. We used categorical variables as potential moderators with Hedges's analogue to ANOVA (Hedges, 1982; Lipsey & Wilson, 2001). Moderator analyses were only conducted if the homogeneity statistic was significant and if an individual meta-analysis of a specific linguistic cue contained enough hypothesis tests to avoid empty cell sizes and to increase power. Moderator analyses were only conducted without outliers to prevent biased results. To clarify potential confounds between moderator variables, we calculated their intercorrelations as well as all two-way and three-way cross-tabulations for each variable combination, to avoid empty or low frequency cells. As a consequence, only moderator analyses for $k \geq 13$ hypothesis tests are reported.

Computer-software for calculations. For computing individual effect sizes, weights and confidence intervals, formulae were programmed in Microsoft Office Excel (2011) spreadsheets by the first and fourth author. Calculations of meta-analyses and outlier analyses were conducted using Excel spreadsheets programmed by the fourth author and cross-validated using Lipsey and Wilson's (2001) SPSS macros (Wilson, 2002). Moderator analyses were also conducted using these macros.

Results and Discussion

Study Characteristics

We included $k = 44$ independent studies or data sets (see Appendix C for all individual coding decisions), dated between 2002 and February 2012. Most studies were published ($k = 27$), 11 were conference presentations (poster or paper), and the rest were 4 Dissertations, 1 Master's Thesis, and 1 submitted manuscript.

Computer program. More than half of the studies (58.1%) used LIWC (2001 or 2007), 23.3% used other general programs, and 18.6% applied a program specifically developed to detect deception. Three studies, where the type of program was not specifically described or labeled (e.g., "automated analysis method", "natural language processing tool", "message analyzing software"), were categorized under other general programs.

Senders. There were a total of 3,780 senders ($k = 43$) with an average of 87.91 (SD = 19.60, Mdn = 53) senders per study, ranging from eight to 800. Information about senders' gender was provided in 30 studies, with more male than female participants in total (N_{male} = 1,254; N_{female} = 895), and on average per study (M_{male} = 41.80; SD_{male} = 9.22; M_{female} = 29.83; SD_{female} = 5.76). Exact information about senders' age was reported in only 29.5% of the studies. Across all age groups, senders' mean age was 19.33 years (SD = 8.45), ranging from 4 to 58 years. The mean age of N = 1,015 adults only was 24.17 (SD = 4.11) with a range of 17 to 58 years, whereas the mean age of N = 218 children ($k = 4$) was 8.45 years (SD = 1.57), ranging from 4 to 14 years.

Accounts. There were a total of 11,680 (N_{truth} = 5,650, N_{lie} = 6,030) accounts originally. However, six studies contained an extremely large number of accounts, ranging from N = 608 (Schafer, 2007, Experiment 1) to N = 3,162 (Derrick et al., 2012), with a mean of 1295.17 accounts (SD = 948.98). In the other 38 studies, the mean was M = 102.87 (SD = 73.17), ranging from N = 13 (Ali & Levine, 2008, confessions) to N = 322 (Cooper, 2008).

Therefore, we decided to adjust the number of total accounts for these six studies to $N = 500$ ($n_{\text{truth}} = 250$, $n_{\text{lie}} = 250$) to avoid extra-ordinary high weights. Consequently, the final average number of accounts per study was $M = 157.02$ ($SD = 153.66$, $Mdn = 103$), with $M = 82.02$ ($SD = 80.71$) for truths and $M = 75.00$ ($SD = 76.68$) for lies. All accounts were provided in English except for four studies (two Spanish, one Dutch, one Arabic).

Preparation. Only eight studies provided information about how long senders had time to prepare their accounts. In four of these, senders had no opportunity, for the other four studies, senders had on average 1.31 minutes ($SD = 0.71$; range: 1 to 5 minutes) to prepare.

Theoretical background. Twelve studies referred to Newman et al.'s (2003) explanations ("LIWC approach") to predict the outcome of specific linguistic cues, three used Interpersonal Deception Theory (IDT, Buller & Burgoon, 1996), two reality monitoring (RM; Sporer, 2004), and 12 a combination of IDT and reality monitoring. Twelve additional studies referred to other theoretical backgrounds, for example, Media Richness Theory (Daft & Lengel, 1986), or Verbal Immediacy (Mehrabian & Wiener, 1966), and three studies did not mention any theory at all. A-priori selections of linguistic cues were made for 37 studies while seven reported only significant findings.

Interpretation of Effect Sizes

As a rule of thumb, Cohen (1988) classified the effect size d into three categories, with $d = 0.20$ as small, $d = 0.50$ as medium and $d = 0.80$ as large effect sizes. However, in meta-analyses about cues to deception, effect sizes are often much smaller (DePaulo et al., 2003: $Mdn\ g_u = 0.10$; similarly low for Sporer & Schwandt, 2006, 2007). Richard, Bond, and Stokes-Zota (2003) examined 322 meta-analyses in social psychology and provided an empirically based effect size distribution that might serve as a good comparison for our results (cf. Sporer & Cohn, 2011). It should be noted that in DePaulo et al.'s (2003) meta-analysis positive effect sizes refer to stronger or more frequent cues in lies.

Research Questions

In this section, we present results for 50 linguistic cues to deception grouped according to six research questions (see Table 1). The weighted average g_u, with the 95% confidence interval (CI), is reported for all analyses. Recall that positive effect sizes denote stronger presence in true accounts (similarly to Sporer & Schwandt, 2006, 2007, but contrary to DePaulo et al., 2003). A data file with all dependent and predictor variables coded is available in supplemental online materials.

Research Question 1: Do Liars Experience Greater Cognitive Load?

(a) Are liars' accounts shorter in terms of *number of words* (cue 01), *number of sentences* (cue 07), and *average sentence length* (cue 08)? As expected, liars used fewer words than truth-tellers (*word quantity*, 0.24 [0.19, 0.29]), with g_us ranging from -1.25 to 1.43, but no shorter sentences than truth-tellers (*average sentence length*, 0.05 [-0.03, 0.13]). Contrary to our prediction, liars used more *sentences* than truth-tellers (-0.33 [-0.44, -0.21]), although the distribution of effect sizes was also quite heterogeneous. The effect size for *sentence quantity* was derived from a small subset of nine studies compared to 42 studies serving data for *word quantity*. Therefore, *word quantity* is a more precise estimate for statement length.

Note that DePaulo et al. (2003) did not examine number of words per se but only *response length* defined as "length or duration" (cue 01, d = -0.03, k = 49, *ns*), or as *talking time* (cue 02, d = -0.35, k = 4, p < .05). Sporer and Schwandt (2006) found no reliable associations for number of words (d = -0.018, k = 8), nor for message duration (d = -0.078, k = 23). These differences in findings may be due to the stimulus accounts used. More recent studies analyzing verbal content cues to deception sometimes do (e.g., Ansarra, Colwell, Hiscock-Anisman, Hines, Fleck, Cole, & Belarde, 2011) and sometimes do not find (e.g., Leal, Vrij, Warmelink, Vernham, & Fisher, 2013) differences between liars' and truth-tellers' length of accounts operationalized by the number of words.

(b) Are deceptive accounts less elaborated in terms of *content word diversity* (cue 02), *type-token ratio* (cue 03), or word length cues (cues 04, 05)? Indeed, liars used *fewer diverse content words* (0.48 [0.34, 0.61]) and distinct words (*type-token ratio*: 0.14 [0.07, 0.21]) than truth-tellers. These findings could be attributed to liars' increased cognitive load and reduced working memory capacity (relative to truth-tellers), which in turn is associated with a limitation of creative word production in speaking or writing. These findings also favor a cognitive over a social psychological explanation, as it is unlikely that liars strategically use fewer diverse content and distinct words. However, the prediction that liars would provide shorter words was not supported (see Table 1). Presumably, the number of distinct words and word diversity indices are more sensitive to cognitive load and working memory capacity than word length.

(c) Are deceptive accounts less complex than true accounts, as indicated by fewer *verbs* (cue 06), *causation* (cue 09) and *exclusive words* (cue 10)? Liars indeed used fewer *exclusive words* like *but*, *except*, or *without*, than truth-tellers (0.24 [0.17, 0.31]). Using few exclusive words results in simpler stories (Newman et al., 2003). Liars may resort to telling simple stories because their cognitive system is more taxed than that of truth tellers. Our predictions that liars would use fewer words assigning a cause to his or her behavior (*causation*), or use fewer *verbs* than truth-tellers, were not confirmed (Table 1).

(d) Do liars commit more *writing errors* (cue 11) than truth-tellers? No support was found for this hypothesis with or without two outliers (Lee, Walker & Odom, 2009; Zhou & Zhang, 2004). This can be reconciled with DePaulo's self-presentational perspective (DePaulo, 1992; DePaulo et al., 2003). Liars might be more self-aware and deliberate than truth-tellers; hence, they may edit their typing errors. Derrick et al. (2012) showed that liars were significantly more likely to edit their words on the keyboard (e.g., in using the backspace and delete button) than truth-tellers (-0.12 [-0.19; -0.05]). Whether or not their edits were aimed at correcting explicit typing errors or not, was not investigated and should be examined more closely. In six of the ten studies exploring *writing errors*, participants typed their stories on a computer keyboard; unfortunately, they did not measure editing behavior (with the exception of Derrick et al., 2012).

Research Question 2: Are Liars Less Certain Than Truth-Tellers?

Effects for *certainty* and *modal verbs* were not significant. The difference between DePaulo et al.'s (2003) findings (who found liars to appear more verbally and vocally uncertain: cue 31, $k = 10$, $d = 0.30$, $p < .05$) and ours could be due to different operationalizations. Whereas we included studies that automatically counted words expressing *certainty*, DePaulo et al. considered the *subjective* impression of uncertainty ("the speaker seems uncertain, insecure, (...)", p. 114). The opposing findings suggest that (a) there is a difference between objective and subjective assessments of (un)certainty, and/or (b) liars may nonverbally give the impression of being uncertain without using fewer certainty words than truth-tellers.

Contrary to our prediction, deceptive accounts contained slightly *fewer tentative words* (such as *may*, *seem*, *perhaps*) than truthful accounts (0.13 [0.06, 0.20] for an exception, see ten Brinke & Porter, 2012). A reason for this unexpected finding could be that liars think that tentative expressions diminish their credibility and therefore try to avoid them, although we are not aware of any empirical evidence that liars pursue this strategy to appear more credible. Note that DePaulo et al. (2003) also reported less "*tentative constructions*" (cue 30, $k = 3$, $d = -0.16$, *ns*) in lies. A different explanation for this finding could be derived from the literature on credibility assessment (e.g., Steller & Köhnken, 1989). The underlying assumption is that due to their motivation to appear credible, liars (here: alleged victims of sexual abuse) would probably not correct themselves spontaneously, admit a lack of memory or raise doubts about their own statement. These criteria relate to uncertainty or tentative words to the extent that liars try to hide any kind of deficiencies or ambiguities in their statement in order to appear or stay credible (Sporer, 2004). Especially the criterion "admitting lack of memory" is less often expressed by liars than truth-tellers (DePaulo et al., 2003, cue 73: $k = 5$, $d = -0.42$, $p < .05$; Vrij, 2005). Research also shows that guilty suspects attempt to be firm in their denial of guilt (Hartwig et al., 2007); this is contrary to showing uncertainty.

Research Question 3a: Do Liars Use More Negations and Negative Emotion Words?

(a) To the extent that liars defend themselves or deny something they have done, do they use more *negation terms* such as *no*, *never*, or *not* (cue 17)? This prediction was supported, with a significant negative effect of -0.15 [-0.22, -0.09] based on 20 studies (but large heterogeneity). Our results contradict Hancock, Curry, Goorha and Woodworth's (2008) view, who considered *negations* as a form of distinction marker (in addition to *exclusive terms*) expected to occur *less* frequently in deceptive accounts, presumably to avoid contradictions by being less specific than truth-tellers.

Our findings concur with those of DePaulo et al. (2003), who found a significant effect for *negative statements and complaints* (cue 52: $d = 0.21$, $k = 9$, $p < .05$) showing that liars use slightly more negative utterances than truth-tellers.

(b) Do liars use more *negative emotion words* in general (cues 18, 18.1), as well as more specific negative-emotion words, such as *anger* (cue 18.2), *anxiety* (cue 18.3), or *sadness* (cue 18.4), than truth-tellers? Contrary to the prediction that people might feel negative emotions while lying (Ekman, 2001; Zuckerman et al., 1981), liars did not use more negative emotion words (cue 18; -0.07 [-0.15, 0.01]). However, the sub-cue *negative emotions only* revealed a small but reliable negative effect (-0.18 [-0.24, -0.12]). The difference between these results can be explained with their different operationalization. Whereas the superordinate category *negative emotions* (cue 18) contained all types of negative emotions (including *anger*, *anxiety*, and *sadness*), cue 18.1 encompassed only a reduced set of negative emotion words (e.g., *hate*, *worthless*, *enemy*).

A more differentiated picture of various negative emotions under investigation emerged when we look at the more specific type of emotion words used. Liars used more *anger* terms than truth-tellers (cue 18.2, -0.27 [-0.35, -0.19]), although no significant differences were found for *anxiety* (cue 18.3) or *sadness* (cue 18.4, see Table 1). Newman et al.'s (2003, p. 672) assertion that "anxiety words are more predictive than overall negative emotion" was not supported. Rather, the present findings indicate that there are differences in words expressing feelings and/or different negative emotions while lying. Liars might not feel anxious or sad but rather feel angry, and this might be manifested in words like *worthless*, or *annoyed*.

Research Question 3b: Do Liars Use Fewer Positive Emotion Words?

Did truth-tellers express more *positive emotion* (cue 19.1) or *positive feeling* (cue 19.2) words than liars? While the effect for *positive emotions only* just missed significance (-0.07 [-0.15, 0.00], overall, there was no support for this prediction (Table 1). DePaulo and colleagues (2003) also did not find a significant effect for being friendly and pleasant (cue 49: d = -0.16, k = 6, *ns*).

Research Question 3c: Do Liars Express More or Less Unspecified Emotion Words?

For 21 studies investigating unspecified *emotion words* (cue 15), liars used more unspecified emotion words than truth-tellers (-0.11 [-0.19, -0.04]). However, liars and truth-tellers did not differ in words expressing *unpleasantness* or *pleasantness* (cue 16, -0.10 [-0.25, 0.06]). DePaulo et al. (2003) also found no significant difference for being "*friendly and pleasant*" (cue 49: d = -0.16, k = 6, *ns*). Conversely, DePaulo et al.'s findings for two other subjectively rated cues associated with pleasantness, namely "*cooperation*" (cue 50: d = -0.66, k = 3, p < .05), and "*facial pleasantness*" (cue 54: d = -0.12, k = 13, p < .05), showed that truth-tellers appeared more pleasant than liars. These differences might indicate that the pleasantness construct tracked by DePaulo et al.'s human-rated cues (subjective impressions) is different from the one operationalized in computer-based studies (objective word count). Alternatively, truth-tellers might only appear more pleasant than liars in their nonverbal or paraverbal behavior, but not in their choice of words.

Research Question 4: Do Liars Distance Themselves More From Events?

(a) Do liars use fewer *first-person pronouns* (cues 21, 22, 23) and more *second-person* (cue 24) *and third-person pronouns* (cue 25) than truth-tellers? Although no significant differences were found for *first-person singular*, or *first-person plural references* (see Table 1), the weighted average effect size for *total first-person pronouns* was significant in the expected direction, that is, liars used fewer *total first-person pronouns* than truth-tellers (0.14 [0.06, 0.22], when the extreme negative effect size found by Brunet, 2009, both conditions: -1.63 [-1.98, -1.29] was excluded).

On the other side of the coin, we predicted *second*-and *third-person pronouns* to occur more often in liars' than truth-tellers' accounts. Our meta-analyses supported this prediction, with a negative g_u = -0.10 (Table 1). The results indicated that liars in general tried to redirect the focus of attention to other people by using more references to their interaction partner(s) (*you*), or to (a) third person(s) (*he*, *she, they*) than truth-tellers.

Overall pronoun use. As researchers seem to be interested in the use of any type of pronouns (*total pronouns*, cue 20), we aggregated all of the pronoun effect sizes. The resulting effect size was not significant (0.06 [-0.02, 0.14]).

(b) Do deceptive accounts contain more *passive voice verbs* (cue 26) and *generalizing terms* (cue 27) than truthful accounts? Although effect sizes for *passive voice* *verbs* varied considerably (see Table 1), all were nonsignificant. This is probably due to small sample sizes or a generally low frequency of occurrence (*floor effect*). *Generalizing terms* had a medium negative effect size (-0.37 [-0.79, 0.05]) that was nevertheless not significant because of the small number of studies and large heterogeneity. Similarly, DePaulo et al. (2003) did not find a significant effect for *generalizing terms* (cue 21: $d = 0.10$, $k = 5$, *ns*).

(c) Do lies include more *past tense* verbs (cue 47) and fewer *present tense* verbs (cue 48) than true accounts? Significant differences were neither found for *past tense verbs* nor for *present tense verbs* (Table 1). A potential reason why the data did not support our predictions could be the way the dependent variable was operationalized. It is important to note that Pillemer et al.'s (1998) hypothesis stated that verb tense *shifts* occur more often in critical parts of experienced (i.e., true) autobiographical events. Here we did not consider verb tense *shifts*, but *absolute number* of present and past tense verbs. Future research could construct a more suitable linguistic cue than counting the number of verbs only.

Research Question 5: Do Liars Use Fewer (Sensory and Contextual) Details?

(a) Do liars use fewer *sensory* and *perceptual* details than truth-tellers? They did, according to our findings for *sensory-perceptual processes only* (cue 28.1), although the average effect size was very small (0.06 [0.00, 0.13]). For the variable *sensory-perceptual processes overall* (cue 28), the effect size was not significant (0.05 [-0.01, 0.12], after two outliers were excluded).

Some support came from the more specialized cue *hearing* (cue 28.4, 0.17 [0.09, 0.25]), showing that liars used fewer words expressing their acoustic impressions (like *listen*, *sound*, or *speak*) than truth-tellers. Indeed, in case of entirely fabricated lies (compared to partially fabricated lies or lies of omission), persons may not experience any audio(visual) impressions at all and do not seem to deliberately include these words in their lies. However, the cues *seeing* (cue 28.2) and *feeling* (cue 28.3) yielded nonsignificant results (see Table 1).

Although DePaulo et al. (2003) also found no significant effects for *sensory information* (cue 05: *d* = -0.17, *k* = 4, *ns*) there have been many new reality monitoring studies we are currently synthesizing in an updated large scale meta-analysis.

(b) Are liars' accounts less contextually embedded than those of truth-tellers, as indicated by fewer *time* and *space* words? No significant effects emerged for *time* (cue 29), *space* (cue 30), or the combination of *spatial and temporal details* (cue 31). Our results for *temporal* and *spatial details* are in line with DePaulo et al.'s (2003) nonsignificant finding for *contextual embedding* (cue 76: *d* = -0.21, *k* = 6, *ns*), though it should be noted that *contextual embedding* goes beyond *temporal* and *spatial details* in that the event has to be connected to everyday occurrences, habits, relationships, and so forth (e.g., Steller & Köhnken, 1989). Again, many newer Criteria-Based Content Analysis and reality monitoring studies found support for this assumption (see Masip et al., 2005; Vrij, 2005, 2008a) but linguistic analyses by computers do not seem to capture them.

(c) Relative to truth-tellers, do liars use fewer descriptive words, such as *prepositions* (cue 32), *numbers* (cue 33), *quantifiers* (cue 34), *modifiers* (adverbs and adjectives, cue 35), and *motion verbs* (cue 36)? The only significant effect size was obtained for *quantifiers* (0.14 [0.02, 0.25]) indicating a slightly lower use of words such as *all, bit, few, less*, among liars. However, this finding was synthesized from four studies only, so we should not make strong conclusions for this cue in general.

Interestingly liars produced more *motion verbs* (such as *walk, go,* or *move*) than truth-tellers (-0.09 [-0.17, -0.01]) after removing the only significant positive effect size (Liu, Hancock, Zhang, Xu, Markowitz, & Bazarova, 2012; 0.38 [0.21, 0.56]), which was found to be an outlier. This finding is contrary to our prediction but is in line with the cognitive load approach (Research Question 1) and Newman et al.'s (2003) assumption that, when constructing a lie, "simple, concrete actions are easier to string together than false evaluations" (p. 667). Therefore, liars, who are cognitively taxed by the act of lying, "should use more *motion verbs* and fewer *exclusive words*" (Newman et al., 2003, p. 667).

Research Question 6: Do Liars Refer Less Often to Cognitive Processes?

As predicted, weighted average effect sizes for both cues (37 and 38) were significantly positive (see Table 1), indicating that liars expressed words relating to their inner thoughts (*insight*) and *cognitive processes* less often than truth-tellers.

Miscellaneous Category

Twenty-nine cues that could not be allocated to any research question were subsumed under the miscellaneous category. As displayed in Appendix D (in supplemental online materials), significant positive effect sizes (without outliers) were obtained for *inhibition*, *humans*, and for three cues expressing biological processes, namely: *biology*, *physical states*, and *eating*. Liars used fewer words from all of these word classes than truth-tellers. In contrast, negative effect sizes for *future tense* and *leisure terms* indicated that these terms occurred more frequently in deceptive than truthful accounts.

Moderator Analyses

Due to the large number of potential moderator analyses for all linguistic cues, we only report significant findings (all Q_Bs were significant at $p < .05$) for both theoretically and methodologically important moderator variables. Specifically, we examined six moderator variables for 25 linguistic cues, with each analysis containing at least 13 studies.[7] Note that the overall number of studies is smaller for the moderator analyses as many studies did not report enough information to code them. Analyses of two additional moderators, experimental design (between- vs. within-participants) and publication status are available in supplemental online materials (Appendices E and F). Also, it must be acknowledged that blocking groups of studies in meta-analyses analogous to ANOVA often introduces confounds (see Pigott, 2012) although we have taken great care to minimize them (see *Method*).

Event type and personal involvement. We hypothesized that larger effect sizes would be found for Research Questions 1, 2, 3a, 3c and 4 if the event was personally relevant to the participant ("First-person experience", k = 21) than in the "Attitude/liking" paradigm (k = 7) or the "Miscellaneous" paradigms (k = 14; see Table 2). First, concerning cognitive load (Research Question 1), event type affected *average sentence length* only. Liars used shorter sentences than truth-tellers when articulating attitudes (0.17), but not under the other two paradigms. Second, regarding negative emotions and negations (Research Question 3a), liars used more *negative emotion words* than truth-tellers only if they had to tell a personally relevant story (-0.37, -0.57), and expressed more *negations* only in miscellaneous paradigms (-0.59). Thus, although liars might experience and express more negative emotions when the topic is personally relevant, they do not necessarily use more negations. Third, as expected, liars also expressed more *unspecified emotions* (Research Question 3c; -0.45) when talking about a personal experience than when having performed other tasks. Fourth, concerning distancing (Research Question 4), liars used fewer *first-person plural pronouns* primarily when describing a video (0.38), but fewer *total first-person pronouns* when talking about attitudes (0.31). This unexpected finding suggests that it may be especially hard for liars to refer to themselves while articulating a false attitude; however, liars may still use self-references while telling a personal event because it is common (in the English language) to refer to oneself as the actor. Also, it would be hard to avoid self-references when telling a story with oneself as the acting person, even when lying.

Liars used more *total second-person pronouns* only when talking about attitudes (-0.18), and more *total third-person pronouns* in all kinds of events except the miscellaneous paradigms. In general, thus, the predicted differences for Distancing (Research Question 4) between liars and truth-tellers appear enhanced in the attitude/liking paradigm--compared to the other two paradigms.

Emotional valence. We predicted effects for cues under Research Questions 1 to 4 to be larger for negative (k = 18) than for neutral events or themes (k = 17; see Table 3).[8] Indeed, liars used fewer *words* (Research Question 1--Cognitive Load; 0.54) only when the event was negative. In terms of negative emotions (Research Question 3a), liars also used considerably more *negations* (-0.42) and *negative emotions* (-0.39, -0.65) than truth-tellers, most notably when the event was negative. This supported the notion that telling a negatively toned lie might be accompanied by negative emotions.

However, contrary to our predictions regarding the cognitive load cues (Research Question 1), differences between lies and truths for *type-token ratio* (0.32) and *exclusive words* (0.47) were larger when telling a neutral event rather than a negative event. Perhaps truth-tellers reporting a negative event are as emotionally involved as liars. This may imply using less elaborate language (compared to neutral events), which would explain the lack of difference in *type-token ratio* between liars and truth-tellers. Finally, no difference in the use of *unspecified emotions* (Research Question 3c) was found between neutral and negative topics: Liars used more *emotion words* overall for both (-0.54; -0.45).

Regarding distancing (Research Question 4), somewhat contradictory findings occurred for self-references. When telling neutral events, liars used more *first-person singular* pronouns (-0.25) but fewer *total first-person* pronouns (0.22) than truth-tellers. Also, when telling negative events liars used fewer *first-person singular* pronouns than truth-tellers (0.27) but about the same amount of *total first-person* pronouns (-0.13). These findings clearly show that (a) differences exist between liars and truth-tellers in terms of referring solely to oneself or to oneself in addition to one's group, and (b) these differences depend on the valence of the event. If we think about examples of wrongdoing as typical negative events, it perfectly makes sense to distribute responsibility to “we” (or “you and me”, “they and me”) than to take it on one’s own shoulders (“I”). Finally, liars expressed more *total second-person pronouns* only when the event was neutral (-0.40).

Intensity of interaction. We predicted that the higher the interaction level, the larger the effect sizes would be (Table 4). Indeed, effect sizes for *word count* (Research Question 1--Cognitive Load; 0.69), *negative emotions* (Research Question 3a; -0.48, -0.79), *unspecified emotions* (Research Question 3c; -0.63), and *first-person singular pronouns* (Research Question 4--Distancing; 0.34) were largest in person to person interactions. Note also that for computer-mediated communication the direction of effect (-0.41) reversed compared to other conditions. Furthermore, in the interview condition (which was considered as the second intense interaction category), effect sizes for *word count*, *exclusive words*, and *negative emotions only*, were in the expected direction. Interestingly, when no interaction took place, *liars* used significantly more *first-person singular pronouns* (-0.22) and *total third-person pronouns* than truth-tellers (-0.31).

Together, this evidence suggests that some verbal differences between liars and truth-tellers manifest themselves most when a bidirectional interaction between two persons--not only a one-way interview--took place.

Motivation. In support for our hypotheses, larger effects occurred for not-motivated liars, who used fewer *words* (Research Question 1--Cognitive Load) than truth-tellers (0.47), compared to moderately (0.19), or highly motivated liars (0.18; see Table 5). Also, liars used fewer *temporal details* (Research Question 5 regarding details) only when no motivation was induced (0.20). These findings support the notion that highly motivated liars are more successful than unmotivated liars in controlling their verbal behavior (at least in terms of number of *words* and *temporal details*). Note that liars seem to be less able to control their paraverbal behavior under high motivation (e.g., for pitch, response latency; Sporer &

Schwandt, 2006) nor their visual nonverbal behavior (e.g., for eye contact; DePaulo et al., 2003).

Other linguistic cues under various research questions showed findings contrary to our hypothesis (see Table 5): (a) only highly motivated liars used fewer *different words* (*type-token ratio*: 0.67); (b) only moderately motivated liars built slightly shorter sentences (*average sentence length*: 0.15) than truth-tellers; (c) liars expressed more *negative emotional* words than truth-tellers only when they were highly (-0.56, -1.03) or not motivated (-0.20, -0.19); (d) liars expressed more *unspecified emotions* (-0.53) than truth-tellers when highly motivated; and (e) both highly (0.21, 0.25) and moderately motivated (both 0.12) liars reported fewer *sensory-perceptual processes* than truth-tellers, whereas not motivated liars tended to refer *more* often to these processes than truth-tellers (-0.22, -0.29).

Taken together, these results show a mixed picture. Our prediction that highly motivated liars would control their verbal behavior better than less motivated liars was confirmed for only two cues. However, our findings should not be over-interpreted because the number of studies with highly motivated participants was very small--calling for more research with highly motivated liars.

Production mode. Moderator analyses showed mixed findings (see Table 6). Liars used fewer *words* (Research Question 1--Cognitive Load) than truth-tellers under all production modes, though effects were larger for handwritten texts (0.33) and for transcripts from spoken accounts (0.26) than for typed texts (0.10). It seems that storytellers may use the opportunity to edit their accounts when typing, thus reducing the number of errors. More direct evidence for this point comes from a study by Derrick et al. (2012) who developed a specialized computer applet that clandestinely recorded edits and revisions during real time synchronous communication between a computer interviewer and senders. They found that when deceiving, people were significantly more likely to take longer and perform a greater number of edits to their responses (more frequently using the delete and backspace keys) than when telling the truth. To the extent that deceptive individuals are more likely to engage in such editing, differences in writing errors between true and false statements may be obscured. This might explain why the effect for number of *typed words* is smaller than for number of handwritten or spoken words.

In line with our hypothesis concerning details (Research Question 5), liars expressed fewer *sensory-perceptual words* than truth-tellers only when writing their accounts by hand (0.33, 0.34), whereas liars used fewer *spatial details* than truth-tellers only in typed accounts (0.13). Contrary to our hypothesis (but in line with Newman et al.'s (2003) assumption), liars used more *motion verbs* than truth-tellers when handwriting (-0.28) or speaking (-0.16), but not when typing them (0.00).

Our hypothesis concerning negative emotions (Research Question 3a) was not supported: Liars expressed more *negations* and *negative emotions* than truth-tellers when handwriting (-0.60 -0.28, respectively) rather than when speaking or typing. A potential reason for the larger effect in the handwriting condition could be that a writer might take more time to re-experience a negative emotion linked to the process of lying (see Ekman, 1988). Also, the special advantage to edit *typed* words could be a reason why the difference between liars and truth-tellers disappeared under this condition. Interestingly, regarding unspecified emotions (Research Question 3c), liars' spoken messages--compared to truth-tellers'--showed no differences (-0.04) in *unspecified emotion words* but more when typing (-0.44) or handwriting (-0.25).

In conclusion, the question of how the mode of production affects the language of lying is not sufficiently answered. Again, other moderators such as interaction type or motivation may be confounded in these analyses. The pattern of findings that typed accounts showed smallest effects also converges with the finding that computer-mediated communication showed smallest effects (Table 4 above). Future studies should investigate interaction intensity and production mode in more detail, perhaps controlling for language proficiency and typing skill.

Computer program. The hypothesis that effects would be larger if statements were analyzed with programs specifically designed to detect deception ($k = 8$) rather than with LIWC ($k = 26$) or general programs ($k = 11$) was only confirmed for *first-person plural pronouns* (-0.31; see Table 7). Specific programs were more sensitive than LIWC or other general programs to differences in *first-person plural* pronouns, finding more of these words among liars than among truth-tellers. Contrary to our hypothesis, four linguistic cues were found to have larger effects if LIWC or other general programs were used than with more specific programs. The direction of the effect for word quantity was even reversed if specific lie detection software was used. A parsimonious explanation may be that these specific programs were developed and used for different types of accounts. It also demonstrates that the validity of linguistic cues to deception depends on the kind of program used. However, this conclusion is limited by the fact that we had to exclude quite a few studies using specialized software as these did not contain sufficient information to calculate effect sizes. Journal editors and grant agencies should emphasize completeness of data reporting including effect sizes (APA, 2008).

Publication bias. The correlation between sample sizes (number of accounts) and the absolute value of all effect sizes (excluding extremely large samples to avoid skewed distributions) was $r(904) = -.11$, $p < .001$. This negative correlation could be due to a publication bias, that is, a tendency for significant findings to be more likely to be published than unpublished (Levine et al., 2009). On the other hand, our moderator analyses showed that for 7 of 12 cues, for which there was a significant difference between published and unpublished studies, effects were actually greater in *unpublished* studies (see Appendix E in supplemental online materials). Thus, publication bias is unlikely to be a threat to the validity of our conclusions.

General Discussion

Setting some of the exceptions discussed under the moderator analyses aside we venture some take home message to our research questions, taking also rival theoretical approaches into consideration.

Research Question 1--Cognitive Load. Taken together, the notion that liars experience greater cognitive load was mainly supported. As predicted from the working memory model and the cognitive load approach, lies were shorter (fewer words and fewer sentences), less elaborated (fewer different words), and less complex (fewer exclusive terms) than true stories. Even if liars were to strategically withhold information that could give them away, doing so would heighten their working memory burden, thus indirectly also supporting the cognitive load approach.

Research Question 2--Certainty. Because only three cues were investigated here and they yielded contradictory results, this question could hardly be answered. In general, the prediction for this research question that liars look linguistically less certain than truth-tellers due to a lack of personal investment or feelings of ambiguity or guilt was not supported. Contrary to our prediction, truth-tellers used more tentative words than liars.

Research Question 3a--Negative Emotions. Altogether, the prediction that liars express more negative emotion words and defend themselves to a greater extent than truth-tellers due to the experience of negative emotions when lying was corroborated. More specifically, liars expressed more terms of anger (rather than other negative emotions like anxiety and sadness) and denied accusations more often than truth-tellers.

Research Question 3b--Positive Emotions. Our assumptions based on the fading-affect bias that truth-tellers express more positive emotions than liars was not supported.

While this result may be dependent on the type of lie being told it does run counter to our assumptions from the autobiographical memory literature (as well as Criteria-Based Content Analysis and reality monitoring research): Taken together, it appears wise to differentiate specific emotions and feelings and separate them according to their valence.

Research Question 3c--Unspecified Emotions. In general, liars expressed more unspecified emotions (i.e., negative and positive emotions undifferentiated) than truth-tellers. Given the results for different types of negative emotions, linguistic researchers should revisit their analyses to separate different types of emotions.

Research Question 4--Distancing. As expected, liars distanced themselves from events more than truth-tellers by using fewer self-references (total first-person) and more other-references (total second- and total third-person). On the other hand, liars and truth-tellers did not differ in terms of generalizing terms, use of passive voice or verb tenses.

Research Question 5--Details. Overall, the reality monitoring approach was only partially supported. We only found small effects for some cues (sensory-perceptual processes only, particularly when motivation is high or the account is handwritten, hearing words and quantifiers) but null-findings for most other cues. In their review on international reality monitoring research, Masip et al. (2005) concluded that visual and auditory details, contextual and temporal information were the most discriminative criteria. The discrepancies may either be due to the fact that the reality monitoring criteria cannot easily be captured by word-counting programs like LIWC, or the fact that the LIWC categories were not created on the basis of reality monitoring theory (see Vrij, Mann, Kristen, & Fisher, 2007). Coding reality monitoring criteria and indicators involves much more than mere word counting, and only well-trained human raters, who also take the context of specific words or sentences, as well as the background or motivation of a statement into account, can do it.

Research Question 6--Cognitions. We found that truth-tellers used more words indicating cognitive processes than liars. The findings support our predictions from autobiographical memory theory that persons refer more often to retrieval processes, supporting memories, and cognitive operations when talking about true events but contradicts the assumption of many reality monitoring deception researchers who postulate the opposite (e.g. Vrij, 2008a).

Limitations

Several limitations restrict the generalizability of our findings. First, we had to exclude more than 50 studies for different reasons (see Appendix G in supplemental online materials). Most of these studies did not provide sufficient statistical data, or calculated linguistic patterns in a way not suitable for our analysis (e.g., Keila & Skillicorn, 2005). While we are grateful to all authors who provided us with additional data, journal editors should emphasize the reporting of all results, not just significant ones, along with effect sizes like Cohen's *d*.

Second, we were able to find significant effects for many linguistic cues (see Table 1). These effects were generally very small according to Cohen (1988), but not much smaller than those for nonverbal and paraverbal cues meta-analyzed by DePaulo et al. (2003) and Sporer and Schwandt (2006, 2007). However, even if all cues had been in the predicted direction, the mean $g_u = 0.26$ is rather disappointing compared to mean effect sizes in the social psychological literature (Richard et al., 2003: mean $g_u = 0.43$; $r = 0.21$, $SD = 0.15$).

Third, for those linguistic cues where effect size distributions were quite heterogeneous, and sensitive to moderator variables, general conclusions can only be very tentative. Specific circumstances of individual studies documented in Appendix C should be considered for specific types of lies, topics, paradigms or production modes.

Fourth, most findings were only available for the English language. As Newman et al. (2003) discussed, deception may be manifested through different linguistic cues in different languages. For example, Romanic languages do not require the use of specific personal pronouns, because pronouns are already expressed by the verb form (e.g., Masip, Bethencourt, Lucas, Sánchez-San Segundo, & Herrero, 2012). Unfortunately, no moderator analysis could be conducted for language, because only four studies analyzed accounts in languages other than English. Besides language, culture might also make a difference. For example, Taylor, Tomblin, Conchie, and Menacere (2011) found that North African participants used first-person pronouns most frequently when lying, whereas White British participants used them most frequently when telling the truth.

Fifth, differences between children and adults became evident as three out of four studies conducted with children were detected as significant outliers, though not for the same cues. This underscores the need to investigate differences between adults' and children's linguistic cues to deception separately. Further, not all children are equal. Linguistic skills develop during childhood, and this may presumably influence the frequency of some potential linguistic deception markers. Children of different ages may show different linguistic cues to deception.

Despite these limitations, our meta-analyses were the first large effort to quantitatively synthesize research in this area. Therefore, they can be seen as the most accurate estimate to date of linguistic differences between liars and truth-tellers assessed by computer programs.

Conclusions and Implications for Future Research

The main goal of the present meta-analysis was to assess the extent to which computer programs are valid and useful tools to detect deception in verbal accounts. We provided clear operational definitions for each cue, derived from an analysis and integration of definitions from different research domains. We then posed theoretically based hypotheses as to the direction of effects for all cues, as well as concerning potential moderator variables. While not all results could be reported due to space limitations, additional appendices and analyses as well as all our raw data are available as supplemental online materials. Researchers are invited to peruse our rich database for additional analyses. Future research should also look at the intercorrelations between linguistic cues to arrive at a better theoretical understanding.

In addition, future research should consider the context of deceptive vs. truthful utterances. A potential reason why only small to medium effect sizes were found in general could be that most computer programs simply count single words without considering the semantic context. If this suggestion goes beyond what computer programs can do at this time, perhaps the linguistic cues with greater effect sizes (for the respective paradigms) should be weighted more heavily than those with smaller or nonsignificant effects. A recent attempt in this direction was made by Chandramouli, Chen, and Subbalakshmi (2011), who employed several weighting mechanisms. They applied for an international patent for their invention of this weighing mechanism.

Ultimately, researchers should directly compare the performance of computers versus human raters. Since context is relevant in analyzing and judging a statement, human raters' assessments of certain linguistic cues might lead to more pronounced differences than objective computer-based codings. On the other hand, the advantages of computer-based coding should not be overlooked. Humans and computers are best at different skills. Humans are less accurate in manual counting of specific cues or in rendering accurate judgments of complex syntactic relationships, whereas computers cannot provide subjective, gestalt like judgments or capture the meaning or intention of what people are saying (for an example of computer-assisted subjective codings see Sporer, 2012).

Finally, we encourage researchers to further investigate the impact of (a) different interview and interaction conditions, (b) mode of production, (c) types of events, (d) age of sender, and (e) language on linguistic markers of deception. In line with Hancock and Woodworth (2013), we found that linguistic cues to deception are sensitive to contextual factors (see moderator variables). These variables are relevant in applied contexts (forensic, work and organizational settings). Researchers should strive to design experiments containing psychological features analogous to real world deceptive situations to enhance ecological validity (e.g., opportunity for preparation, or high motivation).

In sum, to answer the question whether computer programs are effective lie detectors, our answer must be rather skeptical at this time. The effects were not significant for many of the variables studied or small in magnitude, or moderated by situational variables. Alternative theoretical approaches may find other cues or moderators to be important. At this time, researchers', and particularly practitioners', (unrealistic) dream has yet to come true.

References

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

Abe, N. (2009). The neurobiology of deception: Evidence from neuroimaging and loss-of-function studies. *Current Opinion in Neurology, 22,* 594-600. doi:10.1097/WCO.0b013e328332c3cf

Abe, N. (2011). How the brain shapes deception: An integrated review of the literature. *Neuroscientist, 17,* 560-574. doi:10.1177/1073858410393359

Abe, N., Suzuki M., Mori E., Itoh M., & Fujii, T. (2007). Deceiving others: Distinct neural responses of the prefrontal cortex and amygdala in simple fabrication and deception with social interactions. *Journal of Cognitive Neuroscience, 19,* 287-295. doi:10.1162/jocn.2007.19.2.287

Adams, S. H. (2002). *Communication under stress: Indicators of veracity and deception in written narratives* (Doctoral dissertation). Retrieved from http://scholar.lib.vt.edu/ (URN: etd-04262002-164813).

Adams, S. H., & Jarvis, J. P. (2006). Indicators of veracity and deception: An analysis of written statements made to police. *Speech, Language, and the Law, 13,* 1-21. doi:10.1558/ijsll.v13i1.2

*Ali, M. & Levine, T. (2008). The language of truthful and deceptive denials and confessions. *Communication Reports, 21,* 82–91. doi:10.1080/08934210802381862

*Almela Sanchez-Lafuente, A., Valencia-Garcia, R., & Cantos Gomez, P. (2012). Detectando la mentira en lenguaje escrito [Detecting deception in written language]. *Procesamiento de Lenguaje Natural, 48,* 65-72.

Ansarra, R., Colwell, K., Hiscock-Anisman, C., Hines, A., Fleck, R., Cole, L., & Belarde, D. (2011). Augmenting ACID with affective details to assess credibility. *The European Journal of Psychology Applied to Legal Context, 3,* 141-158.

APA Publications and Communications Board Working Group on Journal Article Reporting Standards (2008). Reporting standards for research in Psychology. Why do we need them? What might they be? *American Psychologist, 63,* 839-851. doi:10.1037/0003-066X.63.9.839

Bachenko, J., Fitzpatrick, E., & Schonwetter, M. (2008, August). Verification and implementation of language-based deception indicators in civil and criminal narratives. *Proceedings of the 22nd International Conference on Computational Linguistics* (pp. 41-48), Manchester.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4,* 417–423. doi:10.1016/S1364-6613(00)01538-2

Baddeley, A. (2006) Working memory: An overview. In Pickering S. (Ed.), *Working memory and education* (pp. 1–31). New York: Academic Press.

Bahrick, H. P. (1996). The relation between reproductive and reconstructive processing of memory content. *Behavioral and Brain Sciences, 19,* 191. doi:10.1017/S0140525X00042151

Bahrick, H. P., Hall, L. K., & Berger, S. A. (1996). Accuracy and distortion in memory for high school grades. *Psychological Science, 7,* 265-271. doi:10.1111/j.1467-9280.1996.tb00372.x

*Bedwell, J. S., Gallagher, S., Whitten, S. N., & Fiore, S. M. (2011). Linguistic correlates of self in deceptive oral autobiographical narratives. *Consciousness and Cognition, 20,* 547-555. doi:10.1016/j.concog.2010.10.001

Bender, H.-U. (1987). *Merkmalskombinationen in Aussagen* [Criteria combinations in eyewitness statements]. Tübingen: J. C. B. Mohr.

Benussi, V. (1914). Die Atmungssymptome der Lüge [Breathing symptoms of lying]. In E. Meumann (Ed.), *Sammlung von Abhandlungen zur Psychologischen Pädagogik* (Vol. 3, pp. 513-542). Leipzig: Engelmann.

Blandón-Gitlin, I., Fenn, I., Masip, J., & Yoo, A. (2014). Cognitive-load approaches to detect deception: Searching for cognitive mechanisms. *Trends in Cognitive Sciences, 18*, 441-444. doi:10.1016/j.tics.2014.05.004

Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10,* 214-234. doi:10.1207/s15327957pspr1003_2

*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. doi:10.1002/acp.1087

Borenstein, M. (2009). Effect sizes for continuous data. In H. Cooper, L. V. Harris, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 221-235). New York: Russell Sage Foundation.

Brown, J. D. (2012). Understanding the better than average effect: Motives (still) matter. *Personality and Social Psychology Bulletin, 38,* 209-219. doi:10.1177/0146167211432763

Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition, 5,* 73-99. doi:10.1016/0010-0277(77)90018-X

Bruner, J. (1990). *Acts of meaning.* Cambridge: Harvard University Press.

*Brunet, M. K. (2009). *Why bullying victims are not believed: Differentiating between children's true and fabricated reports of stressful and non-stressful events* (Master's thesis). Retrieved from https://tspace.library.utoronto.ca/.

Brunet, M. K., Evans, A. D., Talwar, V., Bala, N., Lindsay, R. C., & Lee, K. (2013). How children report true and fabricated stressful and non-stressful events. *Psychiatry, Psychology and Law, 20*, 867-881. doi:10.1080/13218719.2012.750896

Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory, 3,* 203-242.

Buller, D. B., Burgoon, J. K., Busling, A., & Roiger, J. (1996). Testing Interpersonal Deception Theory: The language of interpersonal deception. *Communication Theory, 6,* 268-289. doi:10.1111/j.1468-2885.1996.tb00129.x

Bunn, G. C. (2012). *The truth machine: A social history of lie detection.* Baltimore, MD: The Johns Hopkins University Press.

Burgoon, J., Blair, J. P., Qin, T., & Nunamaker, J. F. (2003). Detecting deception through linguistic analysis. In H. Chen (Ed.), *Lecture Notes in Computer Sciences 2665. Intelligence and Security Informatics* (pp. 91–101). Berlin, Germany: Springer-Verlag. doi:10.1007/3-540-44853-5_7

*Burgoon, J. K., & Qin, T. T. (2006). The dynamic nature of deceptive verbal communication. *Journal of Language and Social Psychology, 25,* 76-96. doi:10.1177/0261927X05284482

*Chen, X. (2010). *Psycho-linguistic forensic analysis of internet text data* (Unpublished doctoral dissertation). Faculty of the Stevens Institute of Technology, Hoboken, New Jersey.

Chandramouli, R., Chen, X., & Subbalakshmi, K. P. (2011). *Psycho-linguistic statistical deception detection from text content* (Report No. WO/2011/085108). Retrieved from: http://patentscope.wipo.int/search/en/WO2011085108

Christ, S. E., Van Essen, D. C., Watson, J. M., Brubaker, L. E., & McDermott, K. B. (2009). The contributions of prefrontal cortex and executive control to deception: Evidence from activation likelihood estimate meta-analysis. *Cerebral Cortex, 19,* 1557-66. doi:10.1093/cercor/bhn189

Churyk, N. T., Lee, C.-C., Clinton, D. (2008, October). Can we detect fraud earlier? A technique called content analysis raises the possibility. *Strategic Finance.* Retrieved from www.thefreelibrary.com

Cialdini, R. B., Borden, R. J., Thorne, A., Walker, M. R., Freeman, S., & Sloan, R. L. (1976). Basking in reflected glory: Three (football) field studies. *Journal of Personality and Social Psychology, 34,* 366-375. doi:10.1037/0022-3514.34.3.366

Cody, M. J., Marston, P. J., & Foster, M. (1984). *Deception: Paralinguistic and verbal leakage.* In R. N. Bostrom (Ed.), *Communication yearbook 8* (pp. 464-490). Beverly Hills, CA: Sage.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20,* 37-46. doi:10.1177/001316446002000104

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

College Board (1976–1977). *Student descriptive questionnaire.* Princeton, NJ: Educational Testing Service.

*Colwell, K., Hiscock, C. K., & Memon, A. (2002). Interviewing techniques and the assessment of statement credibility. *Applied Cognitive Psychology, 16,* 287–300. doi:10.1002/acp.78

Conway, M. A. (1990). *Autobiographical memory. An introduction.* Buckingham: Open University Press.

*Cooper, J. E. (2008). *Using natural language processing to identify the rhetoric of deception in business and competitive intelligence email communications* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database (UMI No. 3374690).

Cooper, H. (Ed.) (2010). *Research synthesis and meta-analysis: A step-by-step approach* (4th ed.). Los Angeles: Sage Publications.

Cunningham, H. (2002). GATE, a General Architecture for Text Engineering. *Computers and the Humanities, 36,* 223–254. doi:10.1023/A:1014348124664

Daft, R. L., & Lengel, R. H. (1986). Organizational information requirements, media richness and structural design. *Management Science, 32,* 554-571. doi:10.1287/mnsc.32.5.554

Debey, E., Verschuere, B., & Crombez, G. (2012). Lying and executive control: An experimental investigation using ego depletion and goal neglect. *Acta Psychologica, 140,* 133-141. doi:10.1016/j.actpsy.2012.03.004

DePaulo, B. M. (1992). Nonverbal behavior and self-presentation. *Psychological Bulletin, 111,* 203-243. doi:10.1037/0033-2909.111.2.203

DePaulo, B. M., & Kirkendol, S. E. (1989). The motivational impairment effect in the communication of deception. In J. C. Yuille (Ed.), *Credibility assessment* (pp. 51-70). Dordrecht, The Netherlands: Kluwer. doi:10.1007/BF00987487

DePaulo, B. M., Lanier, K., & Davis, T. (1983). Detecting the deceit of the motivated liar. *Journal of Personality and Social Psychology, 45,* 1096-1103. doi:10.1037//0022-3514.45.5.1096

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129,* 74-118. doi:10.1037/0033-2909.129.1.74

DePaulo, B. M., & Morris, W. L. (2004). Discerning lies from truths: Behavioural cues to deception and the indirect pathway of intuition. In P. A. Granhag, & L. A. Strömwall

(Eds.), *Deception detection in forensic contexts* (pp. 15-40). Cambridge, England: Cambridge University Press. doi:10.1017/CBO9780511490071

*Derrick, D., Meservy, T., Burgoon, J. & Nunamaker, J. (2012, January). An experimental agent for detecting deceit in chat based communication. *Proceedings of the Rapid Screening Technologies,* Deception Detection and Credibility Assessment Symposium, Grand Wilea, Hawaii. doi:10.1109/HICSS.2003.1173793

Dilmon, R. (2009). Between thinking and speaking: Linguistic tools for detecting a fabrication. *Journal of Pragmatics, 41,* 1152-1170. doi:10.1016/j.pragma.2008.09.032

Dulaney, E. F. (1982). Changes in language behavior as a function of veracity. *Human Communication Research, 9,* 75-82. doi:10.1111/j.1468-2958.1982.tb00684.x

Duran, N. D., Crossley, S. A., Hall, C., McCarthy, P. M., & McNamara, D. S. (2009). Expanding a catalogue of deceptive linguistic features with NLP technologies. *Proceedings of the Twenty-Second International Flairs Conference, Sanibel Island, FL,* 243-248.

*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.doi:10.1017/S0142716410000068

Dzindolet, M. T., & Pierce, L. G. (2004). A computerized text analysis can detect deception. *Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting, New Orleans, Louisiana.* doi:10.1177/154193120404800377

*Dzindolet, M. T., & Pierce, L. G. (2005). Using a linguistic analysis tool to detect deception. *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting, Santa Monica, California.* doi:10.1177/154193120504900374

Ekman, P. (1988). Lying and nonverbal behavior: Theoretical issues and new findings. *Journal of Nonverbal Behavior, 12,* 163–176. doi:10.1007/BF00987486

Ekman, P. (2001). *Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage.* New York: W. W. Norton.

Elkins, A. L. (2011). *Vocalic markers of deception and cognitive dissonance for automated emotion detection systems* (Doctoral dissertation). Retrieved from http://arizona.openrepository.com/arizona.

Elntib, S., Wagstaff, G. H., and Wheatcroft, J. M. (2014). The role of account length in detecting deception in written and orally produces autobiographical accounts using reality monitoring. *Journal of Investigative Psychology and Offender Profiling.* Advance online publication. doi:10.1002/jip.1420

Enos, F. (2009). *Detecting deception in speech* (Doctoral dissertation). Retrieved from http://www.cs.columbia.edu/. (UMI No. 3348430).

Enos, F., Shribergy, E., Graciarena, M., Hirschberg, J., & Stolcke, A. (2007, August). Detecting deception using critical segments. *Proceedings of Interspeech, Antwerpen, Belgium.*

*Evans, A. D., Brunet, M. K., Talwar, V., Bala, N., Lindsay, R. C. L., Lee, K. (2012). The effects of repetition on children's true and false reports. *Psychiatry, Psychology and Law, 19,* 517–529. doi:10.1080/13218719.2011.615808

Fleiss, J. L. (1981). *Statistical methods for rates and proportions* (2nd ed.). New York: John Wiley.

Fornaciari, T., & Poesio, M. (2011, June). Lexical vs. surface features in deceptive language analysis. *Proceedings of the ICAIL 2011 Workshop: Applying Human Language Technology to the Law,* Pittsburgh, USA.

Freud, S. (1905/1959). *Bruchstück einer Hysterieanalyse* [Fragments of an analysis of a case of hysteria]. Collected papers. New York: Basic Books. Reprinted in 1959.

Fuller, C. M. (2008). *High-stakes, real-world deception: An examination of the process of deception and deception detection using linguistic-based cues* (Doctoral dissertation). Retrieved from http://dc.library.okstate.edu/cdm/.

*Fuller, C. M., Biros, D. P., Burgoon, J. K., Adkins, M. Twitchell, D. P. (2006). An analysis of text-based deception detection tools. *Proceedings of the 12th Americas Conference on Information Systems, Acapulco, Mexico.*

Fuller, C. M., Biros, D. P., & Delen, D. (2008, January). Exploration of feature selection and advanced classification models for high-stakes deception detection. *Proceedings of the 41st Hawaii International Conference on System Sciences.* doi:10.1109/HICSS.2008.158

Fuller, C. M., Biros, D. P., & Delen, D. (2011). An investigation of data and text mining methods for real world deception detection. *Expert Systems with Applications, 8392-8398.* doi:10.1016/j.eswa.2011.01.032

Fuller, C. M., Biros, D. P., & Wilson (2009). Decision support for determining veracity via linguistic-based cues. *Decision support systems, 46,* 695-703. doi:10.1016/j.dss.2008.11.001

Gamer, M., Bauermann, T., Stoeter, P., & Vosse, G. (2008). Covariations among fMRI, skin conductance, and behavioral data during processing of concealed information. *Human Brain Mapping, 28,* 1287-1301. doi:10.1002/hbm.20343

Global Deception Research Team (2006). A world of lies. *Journal of Cross-Cultural Psychology, 37,* 60–74. doi:10.1080/14789940412331337353

Gombos, V. A. (2006). The cognition of deception: The role of executive processes in producing lies. *Genetic, Social, and General Psychology Monographs, 132,* 197-214. doi:10.3200/MONO.132.3.197-214

Graciarena, M., Shriberg, E., Stolcke, A., Enos, F., Hirschberg, J., & Kajarekar, S. (2006, May). Combining prosodic, lexical and cepstral systems for deceptive speech detection. *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing,* Toulouse, France. doi:10.1109/ICASSP.2006.1660200

Graesser, A. C., McNamara, D. S., Louwerse, M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavioral Research Methods, Instruments, and Computers, 36*, 193–202. doi:10.3758/BF03195564

Granhag, P. A., Strömwall, L. A., & Olsson, C. (2001, June). *Fact or fiction? Adults's abibility to assess children's veracity.* Paper presented at the 11th European Conference on Psychology and Law, Lisbon, Portugal.

Greenhouse, J. B., & Iyengar, S. (2009). Handling missing data. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 417-433). New York: Russell Sage.

Greenwald, A. G. (1980). The totalitarian ego: Fabrication and revision of personal history. *American Psychologist, 35,* 603-618. doi:10.1037/0003-066X.35.7.603

Grubin, D., & Madsen, L. (2005). Lie detection and the polygraph: A historical review. *Journal of Forensic Psychiatry & Psychology, 16*, 357-369.

Gupta, S. (2007) *Modelling deception detection in text* (Master's thesis). Retrieved from http://qspace.library.queensu.ca/handle/1974/922.

Hancock, J. T., Curry, L. E., Goorha, S., & Woodworth, M. (2004, July). Lies in conversation: An examination of deception using automated linguistic analysis. *Proceedings of the 26th Annual Conference of the Cognitive Science Society, Vancouver, Canada.* doi:10.1.1.87.9371

Hancock, J. T., Curry, L. E., Goorha, S., & Woodworth, M. (2005, January). Automated linguistic analysis of deceptive and truthful synchronous computer-mediated communication. *Proceedings of the 38th Hawaii International Conference on System Sciences,* Hawaii. doi:10.1109/HICSS.2005.111

*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. doi:10.1080/01638530701739181

Hancock, J., & Woodworth, M. (2013). An "eye" for an "I": The challenges and opportunities for spotting credibility in a digital world. In B. S. Cooper. D. Griesel, & M. Ternes (Eds.), *Applied issues in investigative interviewing, eyewitness memory, and credibility assessment* (pp. 325-340). New York: Springer.

Hancock, J. T., Woodworth, M., & Goorha, S. (2010). See no evil: The effect of communication medium and motivation on deception detection. *Group Decision and Negotiation, 19,* 327-336. doi:10.1007/s10726-009-9169-7

Hartwig, M., Granhag, P. A., & Strömwall, L. A. (2007). Guilty and innocent suspects' strategies during police interrogations. *Psychology, Crime & Law, 13,* 213-227. doi:10.1080/10683160600750264

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6,* 107-128. doi:10.3102/10769986006002107

Hedges, L. V. (1982). Estimation of effect size from series of independent experiments. *Psychological Bulletin, 92,* 490-499. doi:10.1037/0033-2909.92.2.490

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* New York: Academic Press.

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21,* 1539-1558. doi:10.1002/sim.1186

Hines, A., Colwell, K., Hiscock-Anisman, C., Garrett, E., Ansarra, R., & Montalvo, L. (2010). Impression management strategies of deceivers and honest reporters in an investigative interview. *European Journal of Psychology Applied to Legal Context, 2,* 73-90.

Hirschberg, J., Beus, S., Brenier, J. M., Enos, F., Friedman, S., Gilman, S., Girand, C., Graciarena, M., Kathol, A., Michaelis, L., Pellom, B., Shriberg, E., & Stolcke, A. (2005, September). Distinguishing deceptive from non-deceptive speech. *Proceedings of the 9th European Conference on Speech Communication and Technology, Interspeech,* 1833-1836, Lisboa, Portugal. doi:10.1.1.59.8634

Horowitz, M. W., & Newman, J. B. (1964). Spoken and written expression: An experimental analysis. *Journal of Abnormal and Social Psychology, 68,* 640-647. doi:10.1037/h0048589

*Humpherys, S. L., Moffitt, K. C., Burns, M. B., Burgoon, J. K., & Felix, W. F. (2011). Identification of fraudulent financial statements using linguistic credibility analysis. *Decision Support Systems, 50,* 585–594. doi:10.1016/j.dss.2010.08.009

*Jensen, M. L., Bessarabova, E., Adame, B., Burgoon, J., & Slowik, S. M. (2011). Deceptive language by innocent and guilty criminal suspects: The influence of dominance, question, and guilt on interview responses. *Journal of Language and Social Psychology, 30,* 357-375. doi:10.1177/0261927X11416201

Jensen, M. L., Burgoon, J. K., & Nunamaker, J. F. (2010). Judging the credibility of information gathered from face-to-face interactions. *ACM Journal of Data and Information Quality, 2,* Article 3. doi:10.1145/1805286.1805289.

Jensen, M. L., Lowry, P. B., Burgoon, J. K., & Nunamaker, J. F. (2010). Technology dominance in complex decision making: The case of aided credibility assessment.

Journal of Management Information Systems, 27, 175–201. doi:10.2753/MIS0742-1222270108

Jensen, M. L., Lowry, P. B., & Jenkins, J. L. (2011). Effects of automated and participative decision support in computer-aided credibility assessment. *Journal of Management Information Systems, 28,* 201–233. doi:10.2753/MIS0742-1222280107

Jensen, M. L., Meservy, T. O., Burgoon, J. K., & Nunamaker, J. F. (2010). Automatic, multimodal evaluation of human interaction. *Group Decision and Negotiation, 19,* 367-389. doi:10.1007/s10726-009-9171-0

Johnson, R., Barnhardt, J., & Xhu, J (2004). The contribution of executive processes to deceptive responding. *Neuropsychologia, 42,* 878-901. doi:10.1016/j.neuropsychologia.2003.12.005

Johnson, M. K., Bush, J. C., & Mitchell, K. J. (1998). Interpersonal reality monitoring: Judging the sources of other people's memories. *Social Cognition, 16,* 199-224. doi:10.1521/soco.1998.16.2.199

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source Monitoring. *Psychological Bulletin, 114,* 3-28. doi:10.1037//0033-2909.114.1.3

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88,* 67-85. doi:10.1037//0033-295X.88.1.67

Johnson, M. K., & Suengas, A. G. (1989). Reality monitoring judgments of other people's memories. *Bulletin of the Psychonomic Society, 27,* 107–110.

Keila, P. S., & Skillicorn, D. B. (2005). *Detecting unusual and deceptive communication in email* (ISSN-0836-0227-.2005-498). External Technical Report, School of Computing, Queen's University. Retrieved from http://research.cs.queensu.ca/. doi:10.1.1.68.3131

Kellogg, R. T. (2007). Are written and spoken recall of text equivalent? *The American Journal of Psychology, 120,* 415-428. Retrieved from http://www.jstor.org/stable/20445412

Knapp, M. L., Hart, R. P., & Dennis H. S. (1974). An exploration of deception as a communication construct. *Communication Research, 1,* 15-29. doi:10.1111/j.1468-2958.1974.tb00250.x

*Koyanagi, J. & Blandón-Gitlin, I. (2011, March). *Analysis of children's deception with the Linguistic Inquiry and Word Count approach.* Poster session presented at the 4th International Congress on Psychology and Law / 2011 Annual Meeting of the American Psychology-Law Society, Miami, Florida.

*Krackow, E. (2010). Narratives distinguish experienced from imagined childhood events. *The American Journal of Psychology, 123,* 71-80. Retrieved from: http://www.jstor.org/stable/10.5406/amerjpsyc.123.1.0071

Kuiken, D. (1981). Nonimmediate language style and inconsistency between private and expressed evaluations. *Journal of Experimental Social Psychology, 17,* 183-196. doi:10.1016/0022-1031(81)90013-5

Lane, J. D., & Wegner, D. M. (1995). The cognitive consequences of secrecy. *Journal of Personality and Social Psychology, 69,* 237-253. doi:10.1037/0022-3514.69.2.237

Leal, S., Vrij, A., Warmelink, L., Vernham, Z., & Fisher, R. P. (2013). You cannot hide your telephone lies: Providing a model statement as an aid to detect deception in insurance telephone calls. *Legal and Criminological Psychology.* Advance online publication. doi:10.1111/lcrp.12017

*Lee, C.-C., Welker, R. B., Odom, M. B. (2009). Features of computer-mediated, text-based messages that support automatable, linguistics-based indicators for deception detection. *Journal of Information Systems, 23,* 5-24. doi:10.2308/jis.2009.23.1.24

Lench, H. C., Flores, S. A., & Bench, S. W. (2011). Discrete emotions predict changes in cognition, judgment, experience, behavior, and physiology: A meta-analysis of experimental emotion elicitations. *Psychological Bulletin, 137,* 834-855. doi:10.1037/a0024244

Levine, T., Asada, K. J., & Carpenter, C. (2009). Sample sizes and effect sizes are negatively correlated in meta-analyses: Evidence and implications of a publication bias against nonsignificant findings. *Communication Monographs, 76*, 286-302. doi:10.1080/03637750903074685

Levine, T. R., & McCornack, S. A. (2001). Behavioral adaptation, confidence, and heuristic-based explanations of the probing effect. *Human Communication Research, 27,* 471-502. doi:10.1111/j.1468-2958.2001.tb00790.x

Levine, T. R., Park, H. S., & McCornack, S. A. (1999). Accuracy in detecting truths and lies: Documenting the "veracity effect". *Communication Monographs, 66,* 125-144. doi:10.1080/03637759909376468

Leuprecht, C. (2011). *Deception in speeches of candidates for public office.* Retrieved from Social Sciences Research Network http://papers.ssrn.com/sol3/papers.cfm?abstract_id=173928. doi:10.2139/ssrn.1739282

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* Thousand Oaks: Sage Publications.

*Liu, X., Hancock, J., Zhang, G., Xu, R., Markowitz, D., & Bazarova, N. (2012, January). Exploring linguistic features for deception detection in unstructured text. *Proceedings of the Rapid Screening Technologies, Deception Detection and Credibility Assessment Symposium, Grand Wilea, Hawaii.*

*Masip, J., Bethencourt, M., Lucas, G., Sánchez-San Segundo, M., & Herrero, C. (2012). Deception detection from written accounts. *Scandinavian Journal of Psychology, 53,* 103-111. doi:10.1111/j.1467-9450.2011.00931.x

Masip, J., Garrido, E., & Herrero, C. (2003). Facial appearance and judgments of credibility: The effects of facial babyishness and age on statement credibility. *Genetic Social and General Psychology Monographs, 129,* 269-311.

Masip, J., & Herrero, C. (2013). 'What would you say if you were guilty?' Suspects' strategies during a hypothetical behavior analysis interview concerning a serious crime. *Applied Cognitive Psychology, 27,* 60-70. doi:10.1002/acp.2872

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,* 99-122. doi:10.1080/10683160410001726356

McNamara, D. S. & Graesser, A. C. (2012). Coh-Metrix: An automated tool for theoretical and applied natural language processing. In P. M. McCarthy & C. Boonthum-Denecke (Eds.), *Applied natural language processing: Identification, investigation, and resolution* (pp. 188-205). Hershey, PA: Information Science Reference. doi: 10.4018/978-1-60960-741-8

Mehrabian, A. (1972). *Nonverbal communication.* Chicago: Aldine Publishing Company.

Mehrabian, A., & Wiener, M. (1966). Non-immediacy between communicator and object of communication in a verbal message: Application to inference of attitudes. *Journal of Consulting Psychology, 30,* 420-425. doi:10.1037/h0023813

Meissner, C. A., & Kassin, S. M. (2002). "He's guilty!": Investigator bias in judgments of truth and deception. *Law and Human Behavior, 5,* 469-480. doi:10.1023/A:1020278620751

Mihalcea, R., & Strappavara, C. (2009, August). The lie detector: Explorations in the automatic recognition of deceptive language. *Proceedings of the 47th Annual Meeting of the Association for Computational Linguistics and the 4th International Joint Conference on Natural Language Processing Conference Short Papers, 309–312,* Suntec, Singapore.

*Morgan, C. A., Colwell, K., & Hazlett, G. A. (2011). Efficacy of forensic statement analysis in distinguishing truthful from deceptive eyewitness accounts of highly stressful events. *Journal of Forensic Sciences, 56,* 1227-1234. doi:10.1111/j.1556-4029.2011.01896.x

*Morgan, C. A., Mishara, A., Christian, J., & Hazlett, G. A. (2008, August). Detecting deception through automated analysis of translated speech: Credibility assessment of Arabic-speaking interviewees. *Journal of Intelligence Community Research and Development, 1-22.*

Morgan, C. A., Rabinowitz, Y., Christian, J., & Hazlett, G. A. (2009, January). Detecting deception in Vietnamese: Efficacy of forensic statement analysis when interviewing via an interpreter. *Journal of Intelligence Community Research and Development, 1-16.*

Morgan, C. A., Steffian, G., Clark, W., Coric, V., & Hazlett, G. A. (2008). *Efficacy of Verbal and global judgment cues in the detection of deception in persons interviewed via an interpreter.* Unpublished manuscript.

Neisser, U. (1982). *Memory observed: Remembering in natural contexts.* San Fransico: W. H. Freeman and Company.

*Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (2003). Lying words: Predicting deception from linguistic style. *Personality and Social Psychology Bulletin, 29,* 665-675. doi:10.1177/0146167203029005010

Nunamaker, J. F., Burgoon, J. K., Twyman, N. W., Proudfoot, J. G., Schuetzler, R., & Giboney, J. S. (2012, January). Establishing a foundation for automated human credibility screening. *IEEE International Conference on Intelligence and Security Informatics.* Retrieved from: http://arizona.openrepository.com/arizona/bitstream/10150/222874/1/azu_etd_12045_sip1_m.pdf

*Ott, M., Choi, Y., Cardie, C., & Hancock, J. T. (2011, June). Finding deceptive opinion spam by any stretch of the imagination. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics in Portland, Oregon.*

O'Sullivan, M. (2003). The fundamental attribution error in detecting deception: The boy-who-cried-wolf effect. *Personality and Social Psychology Bulletin, 29,* 1316-1327. doi:10.1177/0146167203254610

Pennebaker, J. W., & Chew, C. H. (1985). Behavioral inhibition and electrodermal activity during deception. *Journal of Personality and Social Psychology, 49,* 1427-1433. doi:10.1037/0022-3514.49.5.1427

Pennebaker, J. W., Francis, M.E., Booth, R.J. (2001). *Linguistic Inquiry and Word Count (LIWC): LIWC 2001.* Mahwah, NJ: Erlbaum.

Pigott, T. D. (2012). *Advances in meta-analysis*. New York: Springer.

Pillemer, D. B., Desrochers, A. B., & Ebanks, C. M. (1998). Remembering the past in the present: Verb tense shifts in autobiographical memory narratives. In C. P. Thompson, D. J. Herrmann, D. Bruce., J. D. Read, D. G. Payne, & M. P. Toglia (Eds.), *Autobiographical memory: Theoretical and applied perspectives* (pp. 145-162). Malwah, NJ: Erlbaum.

Qin, T., & Burgoon, J. (2007, May). An investigation of heuristics of human judgments in detecting deception and potential implications in countering social engineering. *Proceedings of the IEEE Intelligence and Security Informatics,* 152-159. doi:10.1109/ISI.2007.379548

*Qin, T., Burgoon, J., Blair, J. P., & Nunamaker, J. F. (2005). Modality effects in deception detection and applications in automatic deception detection. *Proceedings of the 38th Hawaii International Conference on System Sciences, Hawaii.* doi:10.1109/HICSS.2005.436

Qin, T., Burgoon, J., & Nunamaker, J. F. (2004). An exploratory study on promising cues in deception detection and application of decision tree. *Proceedings of the 37th Hawaii International Conference on System Sciences.* doi:10.1109/HICSS.2004.1265083

Reinhard, M.-A., Sporer, S. L., Scharmach, M., & Marksteiner, T. (2011). Listening, not watching: Situational familiarity and the ability to detect deception. *Journal of Personality and Social Psychology, 101,* 467-484. doi:10.1037/a0023726

Richard, F. D., Bond, C. F., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology, 7,* 331-363. doi:10.1037/1089-2680.7.4.331

Rogers, T. B., Kuiper, N. A., & Kirker, W. S., (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology, 35,* 677-688. doi:10.1037/0022-3514.35.9.677

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (Eds.). (2005). *Publication bias in meta-analysis: Prevention, assessement and adjustments.* Chichester: John Wiley. doi:10.1002/0470870168.ch1

*Rowe, K. & Blandón-Gitlin, I. (2008, March). *Discriminating true, suggested, and fabricated statements with the Linguistic Inquiry and Word Count approach.* Poster presented at the Annual Meeting of the American Psychology-Law Society, Jacksonville, Florida.

Rubin, V. L., & Conroy, N. (2012, March). Discerning truth from deception: Human judgments of automation efforts. *First Monday, 17,* 3. Retrieved from http://firstmonday.org/htbin/

Sauerland, M., & Sporer, S. L. (2011). Written vs. spoken eyewitness accounts: Does modality of testing matter? *Behavioral Sciences and the Law, 29,* 846-857. doi:10.1002/bsl.1013

*Schafer, J. R. (2007). *Grammatical differences between truthful and deceptive narratives* (Unpublished doctoral dissertation). Fielding Graduate University, Santa Barbara, CA.

Schank, R. & Abelson, R. (1977). *Scripts, plans, goals, and understanding: An inquiry into human knowledge structure.* Hillsdale, NJ: Lawrence Erlbaum Associates.

*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. doi:10.1002/jip.121

Shadish, W. R., & Haddock, C. K. (2009). Combining estimates of effect size. In H. Cooper, L. V. Harris, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 257-277). New York: Russell Sage Foundation.

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 592-604. doi:10.1037//0278-7393.4.6.592

Sporer, S. L. (1997). The less travelled road to truth: Verbal cues in deception in accounts of fabricated and self-experienced events. *Applied Cognitive Psychology, 11,* 373–397. doi:10.1002/(SICI)1099-0720(199710)11:5<373::AID-ACP461>3.0.CO;2-0

Sporer, S. L. (1998, March). *Detecting deception with the Aberdeen Report Judgment Scales (ARJS): Theoretical development, reliability and validity.* Paper presented at the Biennial Meeting of the American Psychology-Law Society in Redondo Beach, CA.

Sporer, S. L. (2004). Reality monitoring and the detection of deception. In P.-A. Granhag & L. Stromwall (Eds.), *Deception detection in forensic contexts* (pp. 64-102). Cambridge University Press. doi:10.1017/CBO9780511490071

Sporer, S. L. (2008). Lessons from the origins of eyewitness testimony research in Europe. *Applied Cognitive Psychology, 22,* 737-757. doi:10.1002/acp.1479

Sporer, S. L. (2012). *Making the subjective objective? Computer-assisted quantification of qualitative content cues to deception.* In European Chapter of the Association for Computational Linguistics (Eds.), Proceedings of the Workshop on Computational Approaches to Deception Detection (pp. 78-85). Stroudsberg, PA: Association for Computational Linguistics.

Sporer, S. L. (2013, March). *Content-criteria to detect deception: Methodological pitfalls and solutions.* Paper presented at the Annual Meeting of the American Psychology-Law Society, Portland, USA.

Sporer, S. L., & Cohn, L. D. (2011). Meta-analysis. In B. D. Rosenfeld, & S. D. Penrod (Eds.), *Research methods in forensic psychology* (pp. 43-62). New York: Wiley.

Sporer, S. L., & Küpper, B. (1995). Realitaetsueberwachung und die Beurteilung des Wahrheitsgehaltes von Erzaehlungen: Eine experimentelle Studie [Reality monitoring and the judgment of credibility of stories: An experimental investigation]. *Zeitschrift für Sozialpsychologie, 26,* 173-193.

Sporer, S. L., & Schwandt, B. (2006). Paraverbal correlates of deception: A meta-analysis. *Applied Cognitive Psychology, 20,* 421-446. doi:10.1002/acp.1190

Sporer, S. L., & Schwandt, B. (2007). Moderators of nonverbal indicators of deception: A meta-analytic synthesis. *Psychology, Public Policy, and Law, 13,* 1-34. doi:10.1037/1076-8971.13.1.1

Sporer, S. L., & Sharman, S. J. (2006). Should I believe this? Reality monitoring of accounts of self-experienced and invented recent and distant autobiographical events. *Applied Cognitive Psychology, 20,* 985-1001. doi:10.1002/acp1234

Sporer, S. L., & Walther, A. (2006, March). *Truth Detection by Content Cues: General vs. Specific Questions.* Paper presented at the Meeting of the American Psychology-Law Society in Petersburg, FL.

Steller, M., & Köhnken, G. (1989). Criteria-based statement analysis. In D. C. Raskin (Ed.), *Psychological methods for investigation and evidence* (pp. 217-245). New York: Springer-Verlag.

Strömwall, L. A., Hartwig, M., & Granhag, P. A. (2006). To act truthfully: Nonverbal behavior and strategies during a police interrogation. *Psychology, Crime & Law, 12,* 207-219. doi:10.1080/10683160512331331328

*Suckle-Nelson, J. A., Colwell, K., Hiscock-Anisman, C., Florence, S., Youschak, K. E., & Duarte, A. (2010). Assessment Criteria Indicative of Deception (ACID): Replication and gender differences. *The Open Criminology Journal, 3,* 23-30.

Sutton, A. J. (2009). Publication bias. In H. Cooper, L. V. Harris, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 435-452). New York: Russell Sage Foundation.

Tausczik, Y., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology, 29,* 24-54. doi:10.1177/0261927X09351676

Taylor, P. J., Tomblin, S., Conchie, S. M., Menacere, T. (2011, March). *Verbal indicators of deception in some cultures are indicators of truth in others.* Paper presented at the 4th International Congress on Psychology and Law and the Annual Meeting of the American Psychology-Law Society, Miami, USA.

*ten Brinke, L., & Porter, S. (2012). Cry me a river: Identifying the behavioural consequences of extremely high-stakes interpersonal deception. *Law and Human Behavior, 36,* 469-477. doi:10.1037/h0093929

Toma, C. L., & Hancock, J. T. (2010, February). Reading between the lines: Linguistic cues to deception in online dating profiles. *Proceedings of the ACM Conference on Computer Supported Cooperative Work, 5-8, Savannah, USA.* doi:10.1145/1718918.1718921

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. doi:10.1111/j.1460-2466.2011.01619.x

Twitchell, D. P. (2005). *Automated analysis techniques for online conversations with application in deception detection* (Doctoral dissertation). Retrieved from http://arizona.openrepository.com/arizona/handle/10150/194997

Twitchell, D. P., Biros, D. P., Adkins, M., Forsgren, N., Burgoon, J. K., & Nunamaker, J. F. (2006, January). Automated determination of the veracity of interview statements from people of interest to an operational security force. *Proceedings of the 39th Hawaii International Conference on System Sciences.* doi:10.1109/HICSS.2006.70

Twitchell, D. P., Nunamaker, J. F., & Burgoon, J. K. (2004). Using speech act profiling for deception detection. In H. Chen (Ed.), *Lecture Notes in Computer Sciences 3073. Intelligence and Security Informatics* (pp. 403–410). Berlin, Germany: Springer-Verlag.

*Van Swol, L. M., Braun, M. T., & Malhotra, D. (2012). Evidence for the Pinocchio Effect: Linguistic differences between lies, deception by omissions, and truths. *Discourse Processes, 49,* 79-106. doi:10.1080/0163853X.2011.633331

Vrij, A. (2005). Criteria-Based Content Analysis: A qualitative review of the first 37 studies. *Psychology, Public Policy, and Law, 11,* 3-41. doi:10.1037/1076-8971.11.1.3

Vrij, A. (2008a). *Detecting lies and deceit: Pitfalls and opportunities*. Chichester, England: Wiley.

Vrij, A. (2008b). Nonverbal dominance versus verbal accuracy in lie detection. A plea to change police practice. *Criminal Justice and Behavior, 35,* 1323-1336. doi:10.1177/0093854808321530

Vrij, A., Akehurst, L., Soukara, R., & Bull, R. (2004). Detecting deceit via analyses of verbal and nonverbal behavior in adults and children. *Human Communication Research, 30,* 8–41. doi:10.1111/j.1468-2958.2004.tb00723.x

Vrij, A., Edward, K., Roberts, K. P., Bull, R. (2000). Detecting deceit via analysis of verbal and nonverbal behavior. *Journal of Nonverbal Behavior, 24,* 239–263. doi:10.1023/A:1006610329284

Vrij, A., & Granhag, P. A. (2012). Eliciting cues to deception and truth: What matters are the questions asked. *Journal of Applied Research in Memory and Cognition, 1,* 110–117. doi:10.1016/j.jarmac.2012.02.004

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. doi:10.1007/s10979-006-9066-4

Vrij, A., Mann, S., Leal, S., & Granhag, P. A. (2010). Getting into the minds of pairs of liars and truth-tellers: An examination of their strategies. *The Open Criminology Journal, 3,* 17-22. doi:10.2174/1874917801003010017

Wagner, H., & Pease. K. (1976). The verbal communication of inconsistency between attitudes held and attitudes expressed. *Journal of Personality, 44,* 1-15. doi:10.1111/j.1467-6494.1976.tb00580.x

Walczyk, J. J., Harris, L. L., Duck, T. K., & Mulay, D. (2014). A social-cognitive framework for understanding serious lies: Activation-decision-construction-action theory. *New Ideas in Psychology, 34*, 22-36. doi:10.1016/j.newideapsych.2014.03.001

Walczyk, J. J., Igou, F. P., Dixon, A. P., & Tcholakian, T. (2013). Advancing lie detection by inducing cognitive load on liars: A review of relevant theories and techniques guided by lessons from polygraph-based approaches. *Frontiers in Psychology, 4*(14), 1-13. doi:10.3389/fpsyg.2013.00014

Walczyk, J. J., Mahoney, K. T., Doverspike, D., & Griffith-Ross, D. A. (2009). Cognitive lie detection: Response time and consistency of answers as cues to deception. *Journal of Business and Psychology, 24,* 33-49. doi:10.1007/s10869-009-9090-8

Walczyk, J. J., Schwartz, J. P., Clifton, R., Adams, B., Wei, M., & Zha, P. (2005). Lying person to person about life events: A cognitive framework for lie detection. *Personnel Psychology, 58,* 141-170. doi:10.1111/j.1744-6570.2005.00484.x

Walker, W. R. & Skowronski, J. J. (2009). The fading affect bias: But what the hell is it for? *Applied Cognitive Psychology, 23,* 1122–1136. doi:10.1002/acp.1614

Walker, W. R., Skowronski, J. J., & Thompson, C. P. (2003). Life is pleasant and memory helps to keep it that way. *Review of General Psychology, 7,* 203–210. doi:10.1037/1089-2680.7.2.203

Walker, W. R., Vogl, R. J., & Thompson, C. P. (1997). Autobiographical memory: Unpleasantness fades faster than pleasantness over time. *Applied Cognitive Psychology, 11,* 399–413. doi:10.1002/(SICI)1099-0720(199710)11:5<399::AID-ACP462>3.0.CO;2-E

Watson, K. W. (1981). *Oral and written linguistic indices of deception during employment interviews* (Unpublished doctoral dissertation). Graduate Faculty of the Louisiana State University, Baton Rouge, Louisiana, USA.

Wertheimer, M., & Klein, J. (1904). Psychologische Tatbestandsdiagnostik [Psychological assessment of facts about an event]. *Archiv für Kriminologie, Anthropologie und Kriminalistik, 15,* 72–113.

Wiener, M., & Mehrabian, A. (1968). *Language within language: Immediacy, a channel in verbal communication.* New York: Appleton-Century-Crofts.

*Williams, S. M., Talwar, V., Lindsay, R. C., Bala, N. C., & Lee, K. (2012). Is the truth in your words? Distinguishing children's deceptive and truthful statements. *Journal of Criminology, 2014.* doi:10.1155/2014/547519

Wilson, D. B. (2002). *Meta-analysis macros for SAS, SPSS, and Stata.* Retrieved from http://mason.gmu.edu/~dwilsonb/ma.html

*Zhou, L. (2005). An empirical investigation of deception behavior in instant messaging. *IEEE Transactions on Professional Communication, 48,* 147-160. doi:10.1109/TPC.2005.849652

Zhou, L., Booker, Q. E., & Zhang, D. (2002). ROD: Towards rapid ontology development for underdeveloped domains. *Proceedings of the 35th Hawaii International Conference on System Sciences.* doi:10.1109/HICSS.2002.994046

*Zhou, L., Burgoon, J. K., Nunamaker, J. F., & Twitchell, D. (2004). Automating linguistics-based cues for detecting deception in text-based asynchronous computer-mediated communication. *Group Decision and Negotiation, 13,* 81-106. doi:10.1023/B:GRUP.0000011944.62889.6f

Zhou, L., Burgoon, J. K., Twitchell, D. P., Qin, T. T., & Nunamaker, J. F. (2004). A comparison of classification methods for predicting deception in computer-mediated communication. *Journal of Management Information Systems, 20,* 139-165.

Zhou, L., Burgoon, J. K., Zhang, D. S., & Nunamaker, J. F. (2004). Language dominance in interpersonal deception in computer-mediated communication. *Computers in Human Behavior, 20,* 381-402. doi:10.1016/S0747-5632(03)00051-7

Zhou, L., Shi, Y. M., & Zhang, D. S. (2008). A statistical language modeling approach to online deception detection. *IEEE Transactions on Knowledge and Data Engineering, 20,* 1077-1081. doi:10.1109/TKDE.2007.190624

Zhou, L., & Sung, Y.-W. (2008). Cues to deception in online Chinese groups. *Proceedings of the 41st Hawaii International Conference on System Sciences.* doi:10.1109/HICSS.2008.109

Zhou, L., Twitchell, D. P., Qin, T., Burgoon, J. K., & Nunamaker, J. F. (2003). An exploratory study into deception detection in text-based computer-mediated communication. *Proceedings of the 36th Hawaii International Conference on System Sciences.* doi:0.1109/HICSS.2003.1173793

*Zhou, L., & Zhang, D. (2004). Can online behavior unveil deceivers? An exploratory investigation of deception in instant messaging. *Proceedings of the 37th Hawaii International Conference on System Sciences.* doi:10.1109/HICSS.2004.1265079

Zhou, L., & Zhang, D. S. (2006). A comparison of deception behavior in dyad and triadic group decision making in synchronous computer-mediated communication. *Small Group Research, 37,* 140-164. doi:10.1177/1046496405285125

Zhou, L., & Zenebe, A. (2005). Modeling and handling uncertainty in deception detection. *Proceedings of the 38th Hawaii International Conference on System Sciences,* Hawaii. doi:10.1109/HICSS.2005.438

Zhou, L., & Zenebe, A. (2008, April). Representation and reasoning under uncertainty in deception detection: A neuro-fuzzy approach. *IEEE Transactions on Fuzzy Systems, 16,* 2, 442-454. doi:10.1109/TFUZZ.2006.889914

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 14, pp. 1–60). New York: Academic.

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

Zuckerman, M., Koestner, R. E., Colella, M. J., & Alton, A. O. (1984). Anchoring in the detection of deception and leakage. *Journal of Personality and Social Psychology, 47,* 301-311. doi:10.1037/0022-3514.47.2.301

Footnotes

[1] Similar to Vrij (2008a), we use this term to denote theory regarding both emotions or feelings *and* arousal. While differences between these states have been noted, their overlap has also been acknowledged (Zuckerman et al., 1981).

[2] Ekman (2001) noted that a liar may experience joy ("duping delight"). However, the link between this emotion and verbal cues to deception is not clear (Vrij, 2008a). Therefore, we do not consider it further.

[3] Due to empty cells or small cell sizes in each category, we had to merge previously more differentiated categories to broader categories (see Appendix C).

[4] Although we are aware of some potential confounding variables, such as production mode, communication medium, perspective of sender (e.g., actor or observer), or length of interaction, we developed this moderator variable to find subgroups of studies that were similar in terms of the intensity of interaction between sender and another person. Originally, the categories were more sophisticated, but due to small cell sizes, we had to collapse some related categories.

[5] Although we consider word count an important variable, this variable has been investigated in several other meta-analyses (e.g., Sporer & Schwandt, 2006; DePaulo et al. 2003, and Zuckerman & Driver, 1985: combination of *duration* and *number of words*), plus in all the studies that investigated linguistic cues summarized here. Also, many studies on content cues to deception assessed by humans have reported on word count, usually by using a word processor. To review all these studies (likely to be several hundred) where the main focus was not on computer-aided detection of deception would constitute a meta-analysis of its own and is beyond the scope of this paper.

[6] Even when four outliers (with two positive and two negative values) were excluded for *negative emotions only*, the effect remained significant (k = 20, -0.12 [-0.19, -0.04]).

[7] Due to the fairly large number of potential pairwise comparisons for each moderator variable and linguistic cue, we did not calculate these specific comparisons. More differentiated results for homogeneity test statistics between and within groups as well as all other (e.g., nonsignificant) moderator-analytic results can be requested from the first author.

[8] Because purely positive events (k = 3) as well as a combination of positive and negative events (k = 6) were quite rarely used, they were excluded from moderator analyses.

Table 1

Meta-Analyses of Linguistic Cues under Research Questions 1 to 6

Linguistic Cue	Pred. DOE	*k*	*N*	Min g_u	Max g_u	g_u	CI-low	CI-high	*Q*	I^2
Research Question 1: Do Liars Experience Greater Cognitive Load?										
(a) Length of Accounts										
01 Word Quantity[M]	T	42	6,713	-1.25	1.43	**0.24**	0.19,	0.29	**315.85**	87.02
07 Sentence Quantity	T	9	1,334	-1.31	0.28	**-0.33**	-0.44,	-0.21	**104.01**	93.31
08 Average Sentence Length[wO,M]	T	15/16	2,704	-0.37	0.43	0.05	-0.03,	0.13	20.46	31.58
(b) Elaboration of Accounts										
02 Content Word Diversity[wO]	T	7/9	1,076	0.27	1.00	**0.48**	0.34,	0.61	8.13	26.22
03 Type-Token Ratio[M]	T	22	3,589	-1.40	1.09	**0.14**	0.07,	0.21	**171.95**	87.79
04 Six Letter Words	T	10	1,617	-0.28	0.25	-0.05	0.14,	-0.05	5.19	0.00
05 Average Word Length[wO]	T	7/8	954	-0.42	0.69	0.11	-0.03,	0.25	6.85	12.41
(c) Complexity of Accounts										
06 Verb Quantity	T	12	2,356	-1.21	0.44	-0.03	0.11,	-0.60	**78.91**	86.06
09 Causation	T	17	2,773	-0.68	0.25	-0.07	-0.14,	0.01	18.61	14.03

10 Exclusive Words[wO,M]	T	18/20	2,783	-0.24	0.81	**0.24**	0.17,	0.31	25.87	25.66
(d) Errors in Production										
11 Writing Errors[wO]	D	8/10	990	-0.65	0.43	-0.01	-0.15,	0.12	13.36	47.61
Research Question 2: Are Liars Less Certain Than Truth-Tellers?										
12 Tentative Words[wO]	D	19/20	3,145	-1.27	0.36	**0.13**	0.06,	0.20	20.13	10.56
13 Modal Verbs	D	25	3,889	-0.42	0.80	0.00	-0.07,	0.06	32.73	26.62
14 Certainty	T	18	2,823	-0.25	0.94	-0.06	-0.14,	0.01	25.15	32.40
Research Question 3a: Do Liars Use More Negations and Negative Emotion Words?										
17 Negations[M]	D	20	3,659	-0.98	0.53	**-0.15**	-0.22,	-0.09	**155.53**	87.78
18 Negative Emotions[+,wO,M]	D	21/24	2,593	-0.39	0.87	-0.07	-0.15,	0.01	21.03	4.88
18.1 Negative Emotions Only[M]	D	24	3,641	-1.90	0.87	**-0.18**	-0.24,	-0.12	**214.57**	89.28
18.2 Anger	D	12	2,452	-1.32	0.38	**-0.27**	-0.35,	-0.19	**165.03**	93.34
18.3 Anxiety[wO]	D	11/12	1,952	-0.30	0.44	0.07	-0.02,	0.02	13.55	26.19
18.4 Sadness	D	12	2,452	-0.34	0.25	0.04	-0.04,	0.12	13.01	15.46
Research Question 3b: Do Liars Use Fewer Positive Emotion Words?										

19 Positive Emotions and Feelings[+,wO]	T	20/21	2,703	-0.84	0.37	-0.05	-0.12,	0.03	29.59	35.79
19.1 Positive Emotions Only[wO,M]	T	20/21	2,703	-0.84	0.35	-0.07	-0.15,	0.00	27.98	32.08
19.2 Positive Feelings Only	T	9	1,422	-0.47	0.40	0.07	-0.03,	0.18	14.88	46.25
Research Question 3c: Do Liars Express More or Less Unspecified Emotion Words?										
15 Emotions[wO,M]	?	21/25	2,941	-0.63	0.48	**-0.11**	-0.19,	-0.04	28.92	30.85
16 Pleasantness and Unpleasantness	?	6	806	-0.35	0.30	-0.10	-0.25,	0.06	8.43	40.68
Research Question 4: Do Liars Distance Themselves More From Events?										
(a) Personal Pronouns										
21 First-Person Singular[M]	T	22	3,761	-1.00	0.61	-0.06	-0.13,	0.00	**274.38**	92.35
22 First-Person Plural[wO,M]	T	22/25	3,224	-0.72	0.39	0.06	-0.01,	0.13	27.98	24.93
23 Total First-Person[wO,M]	T	22/23	2,541	-0.39	0.57	**0.14**	0.06,	0.22	32.26	34.91
24 Total Second-Person[wO,M]	D	21/23	3,072	-0.61	0.33	**-0.10**	-0.17,	-0.02	28.64	30.16
25 Total Third-Person[wO,M]	D	26/29	3,848	-0.41	0.55	**-0.10**	-0.16,	-0.04	37.20	32.79
20 Total Pronouns[wO,M]	-	18/19	2,460	-0.36	0.65	0.06	-0.02,	0.14	19.32	12.02
(b) Passive Voice and Generalizing Terms										

26 Passive Voice Verbs	D	11	1,221	-0.47	0.49	0.06	-0.06,	0.18	9.52	0.00
27 Generalizing Terms	D	4	93	-1.63	0.44	-0.37	-0.79,	0.05	**15.78**	80.99
(c) Past and Present Tense										
47 Past Tense	D	16	3,047	-0.53	0.41	0.06	-0.01,	0.14	22.67	33.83
48 Present Tense[wO,M]	T	16/17	2,607	-0.51	0.60	0.01	-0.07,	0.09	19.38	22.61
Research Question 5: Do Liars Use Fewer (Sensory and Contextual) Details?										
(a) Sensory and Perceptual Details										
28 Sens.-Perceptual Processes[+,wO,M]	T	25/27	3,957	-0.70	0.70	0.05	-0.01,	0.12	36.00	33.33
28.1 Sens.-Percept. Processes Only[M]	T	27	4,177	-0.90	0.70	**0.06**	0.00,	0.13	**89.26**	70.87
28.2 Seeing[wO]	T	9/11	1,740	-0.17	0.34	0.03	-0.06,	0.13	13.34	40.03
28.3 Feeling[wO]	T	11/12	2,304	-0.49	0.27	-0.03	-0.11,	0.05	15.00	33.31
28.4 Hearing	T	11	2,344	-0.41	0.48	**0.17**	0.09,	0.25	14.15	29.35
(b) Contextual Embedding										
29 Time[wO]	T	23/24	3,296	-1.25	0.53	0.03	-0.04,	0.10	28.95	24.00
30 Space[wO,M]	T	22/24	3,199	-0.36	0.58	-0.04	-0.13,	0.03	31.74	33.74

31 Space & Time	T	5	634	-0.25	0.61	-0.04	-0.19,	0.12	**10.48**	61.84
(c) Descriptive Words										
32 Prepositions	T	14	2,479	-0.55	0.48	0.02	-0.06,	0.10	16.54	21.38
33 Numbers	T	12	2,452	-0.28	0.23	0.05	-0.03,	0.13	9.37	0.00
34 Quantifier	T	4	1,198	0.06	0.22	**0.14**	0.02,	0.25	1.22	0.00
35 Modifier	T	11	1,361	-1.04	0.43	-0.08	-0.20,	0.03	**77.46**	87.09
36 Motion Verbs[wO,M]	T	16/17	2,359	-0.72	0.13	**-0.09**	-0.17,	-0.01	16.84	10.92
Research Question 6: Do Liars Refer Less Often to Cognitive Processes?										
37 Cognitive Processes[wO]	T	18/19	2,915	-0.25	0.36	**0.09**	0.01,	0.16	19.66	13.54
38 Insight[M]	T	15	2,539	-0.41	0.59	**0.13**	0.05,	0.21	**35.65**	60.73

Notes. g_u = effect size Hedges' g_u, positive g_us indicate higher frequencies in true accounts, negative g_us indicate higher frequencies in deceptive accounts; [wO] = without outlier: Results after removal of outliers detected using Hedges and Olkin's (1985) procedure; [M] = Moderator analyses conducted; [+]indicates that the specific linguistic cue is an umbrella term; Pred. DOE = predicted direction of effect; T = occurs more often in true accounts: D = occurs more often in deceptive accounts; k = number of hypothesis tests, where the second value behind the slash indicates the number of hypothesis tests with outliers; N = total number of accounts; *Min* = minimum; *Max* = maximum; Q = homogeneity test statistic; CI = 95% confidence interval; I^2 = descriptive measure of heterogeneity; **values in bold** indicate significance ($p < .05$).

Table 2

Effect Sizes of Linguistic Cues to Deception when Studies used Different Type of Events and Personal Involvement

Linguistic Cue (Research Question, RQ)	*k*	Overall g_u [CI]	k_1	Attitude/Liking Paradigm	k_3	First-Person Experience	k_2	Miscellaneous Paradigms
08 Average Sentence Length[wO] (RQ1)	15	0.04 [-0.04, 0.12]	2	***0.17 [0.05, 0.29]***	8	-0.07 [-0.20, 0.06]	5	-0.07 [-0.17, 0.13]
17 Negations (RQ3a)	20	**-0.15 [-0.22, -0.08]**	7	0.08 [-0.02, 0.18]	7	-0.08 [-0.20, 0.05]	6	***-0.59 [-0.71, -0.47]***
18 Neg. Emotions[+] (RQ3a)	24	**-0.13 [-0.19, -0.06]**	7	0.03 [-0.06, 0.13]	10	***-0.37 [-0.48, -0.25]***	7	-0.10 [-0.24, 0.03]
18.1 Neg. Emotions Only (RQ3a)	24	**-0.17 [-0.24, -0.11]**	7	0.06 [-0.04, 0.15]	10	***-0.57 [-0.69, -0.45]***	7	-0.11 [-0.25, 0.03]
15 Emotions (RQ3c)	24	**-0.21 [-0.27, -0.14]**	7	-0.08 [-0.17, 0.02]	12	***-0.45 [-0.56, -0.34]***	5	-0.02 [-0.19, 0.15]
22 First-Person Plural (RQ4)	24	**0.08 [0.01, 0.14]**	7	0.09 [-0.01, 0.19]	12	-0.08 [-0.18, 0.03]	5	***0.38 [0.23, 0.53]***
23 Total First-Person[wO] (RQ4)	22	**0.14 [0.06, 0.22]**	6	**0.31 [0.19, 0.42]**	11	-0.11 [-0.27, 0.05]	5	0.10 [-0.05, 0.26]
24 Total Second-Person (RQ4)	22	-0.04 [-0.01, 0.03]	7	***-0.18 [-0.27, -0.08]***	10	0.09 [-0.02, 0.20]	5	0.09 [-0.09, 0.26]
25 Total Third-Person (RQ4)	28	**-0.11 [-0.17, -0.05]**	7	**-0.12 [-0.21, -0.02]**	**13**	***-0.22 [-0.32, -0.12]***	8	0.06 [-0.07, 0.18]

Notes. *k* = number of hypothesis tests; g_u = effect size (ES) Hedges' g_u, positive g_us indicate higher frequencies in true accounts, negative g_us indicate higher frequencies in deceptive accounts; CI = 95% confidence interval; [wO] = without Outlier; [+] indicates that the specific linguistic cue is an umbrella term; Neg.= Negative; **bold** ES indicate significant difference from zero; ES in *italics* correspond to the largest difference between liars and truth tellers (in the predicted direction according to the RQ) for a specific cue across the moderator variable levels; this indicates under what level of the moderator variable our hypotheses were most strongly supported.

Table 3

Effect Sizes of Linguistic Cues to Deception when the Emotional Valence of the Event was Neutral versus Negative

Linguistic Cue (Research Question, RQ)	k	Overall g_u [CI]	k_1	Neutral	k_2	Negative
01 Word Quantity (RQ1)	33	**0.25 [0.19, 0.31]**	17	0.04 [-0.03, 0.12]	16	***0.54 [0.45, 0.62]***
03 Type-Token Ratio (RQ1)	20	**0.26 [0.18, 0.33]**	11	***0.32 [0.24, 0.41]***	9	0.04 [-0.12, 0.19]
10 Exclusive Words (RQ1)	14	**0.38 [0.29, 0.46]**	8	***0.47 [0.36, 0.59]***	6	**0.26 [0.11, 0.38]**
17 Negations (RQ3a)	14	**-0.30 [-0.38, -0.21]**	8	**-0.13 [-0.26, 0.01]**	6	***-0.42 [-0.53, -0.31]***
18 Neg. Emotions (RQ3a)	17	**-0.26 [-0.35, -0.18]**	10	**-0.16 [-0.28, -0.05]**	7	***-0.39 [-0.52, -0.26]***
18.1 Neg. Emotions Only^{wO} (RQ3a)	17	**-0.41 [-0.50, -0.32]**	10	**-0.22 [-0.34, -0.10]**	7	***-0.65 [-0.79, -0.52]***
15 Emotions (RQ3c)	18	**-0.50 [-0.58, -0.42]**	8	**-0.54 [-0.65, -0.43]**	10	**-0.45 [-0.57, -0.33]**
21 First-Person Singular^{wO} (RQ4)	15	-0.04 [-0.12, 0.05]	9	**-0.25 [-0.36, -0.13]**	6	***0.27 [0.14, 0.40]***
23 Total First-Person^{wO} (RQ4)	17	**0.10 [0.01, 0.20]**	8	***0.22 [0.10, 0.34]***	9	-0.13 [-0.30, 0.05]
24 Total Second-Person^{wO} (RQ4)	15	**-0.20 [-0.28, -0.12]**	9	***-0.40 [-0.50, -0.29]***	6	0.09 [-0.05, 0.22]

Notes. Please consult notes from Table 2.

Table 4

Effect Sizes of Linguistic Cues to Deception when Studies Applied Different Type of Interaction Levels (between Sender and Receiver)

Linguistic Cue (Research Question, RQ)	k	Overall g_u [CI]	k_1	No Interaction	k_2	Computer-Mediated Communication	k_3	Interview	k_4	Person to Person Interaction
01 Word Quantity (RQ1)	37	**0.22 [0.17, 0.28]**	13	**0.14 [0.07, 0.21]**	5	**-0.41 [-0.65, -0.18]**	16	**0.35 [0.23, 0.46]**	3	***0.69 [0.53, 0.85]***
10 Exclusive Words (RQ1)	19	**0.30 [0.23, 0.36]**	9	***0.37 [0.29, 0.45]***	2	-0.02 [-0.31, 0.26]	6	**0.25 [0.04, 0.46]**	2	**0.17 [0.02, 0.33]**
18 Neg. Emotions[+] (RQ3a)	22	**-0.14 [-0.21, -0.07]**	8	0.03 [-0.07, 0.12]	3	-0.18 [-0.46, 0.10]	7	-0.16 [-0.34, 0.01]	4	***-0.48 [-0.63, -0.34]***
18.1 Neg. Emotions Only (RQ3a)	22	**-0.20 [-0.27, -0.13]**	8	0.05 [-0.05, 0.14]	3	-0.18 [-0.46, 0.10]	7	**-0.18 [-0.36, -0.01]**	4	***-0.79 [-0.94, -0.64]***
15 Emotions (RQ3c)	24	**-0.34 [-0.41, -0.28]**	11	**-0.36 [-0.44, -0.28]**	0		11	-0.04 [-0.19, 0.11]	2	***-0.63 [-0.79, -0.47]***
21 First-Person Singular[wO] (RQ4)	21	-0.06 [-0.12, 0.01]	9	**-0.22 [-0.30, -0.13]**	2	-0.04 [-0.33, 0.24]	7	0.01 [-0.16, 0.19]	3	***0.34 [0.20, 0.49]***
25 Total Third-Person (RQ4)	27	**-0.21 [-0.27, -0.15]**	10	***-0.31 [-0.40, -0.23]***	2	-0.21 [-0.52, 0.09]	12	-0.04 [-0.18, 0.09]	3	-0.09 [-0.23, 0.06]

Notes. Please consult notes from Table 2.

Table 5

Effect Sizes of Linguistic Cues to Deception when Studies Induced Different Levels of Motivation

Linguistic Cue (Research Question, RQ)	*k*	Overall g_u [CI]	k_1	No Motivation	k_2	Low to Medium Motivation	k_3	High Motivation
01 Word Quantity (RQ1)	37	**0.27 [0.21, 0.32]**	11	***0.47 [0.36, 0.57]***	22	**0.19 [0.12, 0.26]**	4	**0.18 [0.04, 0.31]**
03 Type-Token Ratio (RQ1)	19	0.00 [-0.08, 0.08]	4	-0.17 [-0.39, 0.05]	12	**-0.12 [-0.21, -0.02]**	3	***0.67 [0.47, 0.87]***
08 Average Sentence Length[wO] (RQ1)	13	0.08 [-0.01, 0.16]	5	0.09 [-0.09, 0.28]	6	**0.15 [0.04, 0.26]**	2	-0.21 [-0.42, 0.01]
18 Neg. Emotions[+] (RQ3a)	21	**-0.14 [-0.21, -0.07]**	6	**-0.19 [-0.36, -0.03]**	13	-0.01 [-0.10, 0.07]	2	***-0.56 [-0.74, -0.39]***
18.1 Neg. Emotions Only (RQ3a)	21	**-0.20 [-0.27, -0.13]**	6	**-0.20 [-0.37, -0.03]**	13	0.00 [-0.09, 0.09]	2	***-1.03 [-1.21, -0.86]***
15 Emotions (RQ3c)	23	**-0.22 [-0.29, -0.15]**	4	-0.20 [-0.42, 0.01]	15	**-0.10 [-0.19, -0.02]**	4	***-0.53 [-0.66, -0.39]***
28 Sens.-Perc. Processes[+] (RQ5)	25	**0.08 [0.02, 0.15]**	7	**-0.22 [-0.39, -0.06]**	15	**0.12 [0.03, 0.20]**	3	***0.21 [0.07, 0.35]***
28.1 Sens.-Perc. Processes Only (RQ5)	25	**0.08 [0.02, 0.15]**	7	**-0.29 [-0.45, -0.13]**	15	**0.12 [0.03, 0.20]**	3	***0.25 [0.12, 0.38]***
29 Time[wO] (RQ5)	21	0.03 [-0.05, 0.10]	4	***0.20 [0.01, 0.40]***	15	-0.02 [-0.10, 0.07]	2	0.09 [-0.12, 0.30]

Notes. Please consult notes from Table 2.

Table 6

Effect Sizes of Linguistic Cues to Deception when Studies used Different Modes of Producing an Account

Linguistic Cue (Research Question, RQ)	*k*	Overall g_u [CI]	k_1	Handwritten	k_2	Typed	k_3	Spoken
01 Word Quantity (RQ1)	38	**0.19 [0.13, 0.24]**	6	***0.33 [0.21, 0.44]***	14	**0.10 [0.03, 0.17]**	18	**0.26 [0.15, 0.36]**
17 Negations (RQ3a)	19	**-0.19 [-0.26, -0.12]**	5	***-0.60 [-0.72, -0.46]***	5	0.06 [-0.05, 0.17]	9	-0.14 [-0.28, -0.01]
18.1 Neg. Emotions Only (RQ3a)	23	-0.04 [-0.11, 0.03]	3	***-0.28 [-0.51, -0.05]***	8	0.07 [-0.03, 0.16]	12	**-0.14 [-0.25, -0.02]**
15 Emotions (RQ3c)	21	**-0.29 [-0.36, -0.22]**	4	**-0.25 [-0.44, -0.07]**	6	***-0.44 [-0.54, -0.35]***	11	-0.04 [-0.16, 0.08]
28 Sens.-Perc. Processes^{+} (RQ5)	24	0.06 [-0.01, 0.13]	3	***0.33 [0.11, 0.54]***	8	0.01 [-0.08, 0.11]	13	0.06 [-0.06, 0.17]
28.1 Sens.-Perc. Processes Only (RQ5)	24	0.05 [-0.01, 0.12]	3	***0.34 [0.12, 0.56]***	8	0.00 [-0.09, 0.10]	13	0.05 [-0.07, 0.16]
30 Space (RQ5)	22	0.04 [-0.03, 0.11]	3	0.03 [-0.19, 0.24]	6	***0.13 [0.02, 0.23]***	13	-0.06 [-0.17, 0.05]
36 Motion Verbs^{wO} (RQ5)	16	**-0.09 [-0.17, -0.01]**	2	**-0.28 [-0.57, 0.00]**	4	0.00 [-0.11, 0.11]	10	**-0.16 [-0.29, -0.04]**

Notes. Please consult notes from Table 2.

Table 7

Effect Sizes of Linguistic Cues to Deception when Studies used LIWC, a General Program or a Specific Program

Linguistic Cue	*k*	Overall g_u [CI]	k_1	LIWC	k_2	General Program	k_3	Specific Program
01 Word Quantity	41	**0.25 [0.20, 0.30]**	23	**0.28 [0.21, 0.34]**	10	***0.53 [0.43, 0.63]***	8	**-0.19 [-0.30, -0.07]**
15 Emotions	25	**-0.25 [-0.41, -0.28]**	19	**-0.39 [-0.45, -0.32]**	-	-	6	-0.14 [-0.29, 0.02]
17 Negations	20	**-0.15 [-0.22, -0.08]**	17	0.05 [-0.03, 0.12]	3	***-0.82 [-0.96, -0.69]***	-	-
20 Total Pronouns[wO]	18	**0.29 [0.20, 0.38]**	13	**0.38 [0.28, 0.49]**	2	0.06 [-0.17, 0.28]	3	-0.13 [-0.44, 0.17]
			k_1	LIWC + General Program			k_2	Specific Program
22 First-Person Plural	25	-0.04 [-0.10, 0.02]	18	0.01 [-0.05, 0.08]			7	**-0.31 [-0.46, -0.15]**

Notes. Please consult notes from Table 2.

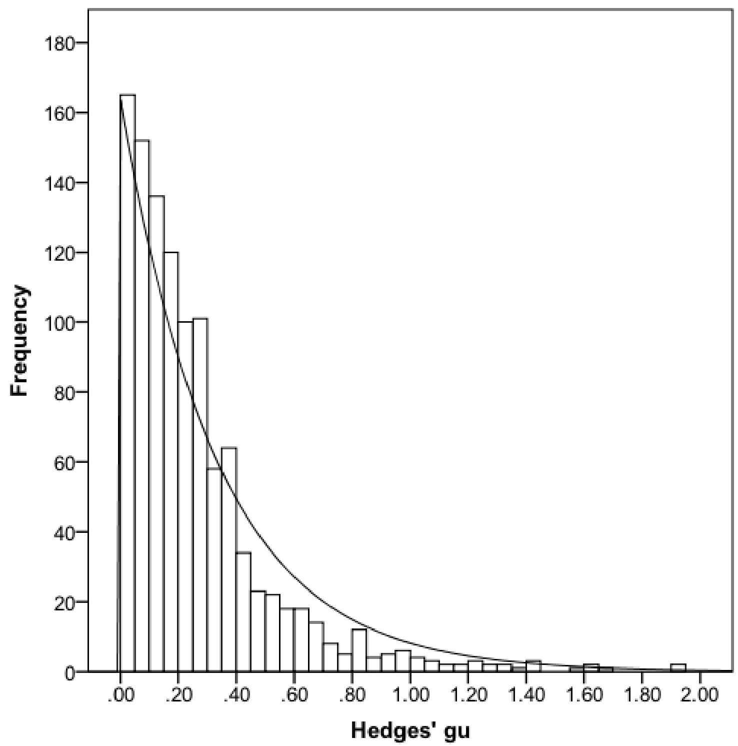

Figure 1. Distribution of all absolute values effect sizes (k = 1,093).

Appendix A

Definition of Linguistic Cues to Deception assigned to Research Questions

	Linguistic Cue	Final Operational Definition
		Research Question 1: Do Liars Experience Greater Cognitive Load?
01*	**Word Quantity** // Word Count // Number of Words // Productivity	Total number of words.
02	**Content Word Diversity** // Diversity // Content Diversity	Total number of *different* content words divided by total number of content words, where content words express lexical meaning.
03	**Type-Token Ratio** // Unique Words // Lexical Diversity // Different Words	% of distinct words divided by total number of words.
04	**Six letter words** // Percentage Words longer than six letters	% of words that are longer than six letters.
05	**Average Word Length** (AWL; Complexity) // Lexical complexity	Total number of letters divided by the total number of words.
06*	**Verb Quantity** // Verb Count	Total number of verbs.
07*	**Sentence Quantity** // Number of Sentences	Total number of sentences.
08	**Average Sentence Length** (Complexity Measure) // Words per Sentence	Total number of words divided by total numbers of sentences.
09	**Causation**	% of words that try to assign a cause to whatever the person is describing (e.g., because, effect, hence).

10	**Exclusive**	% of words that make a distinction what is in a category and what is not (e.g., without, except, but).
11	**Writing Errors** // Typographical error ratio (Informality) // Typo ratio // Misspelled Words	% of writing errors or misspelled words divided by number of words.
	Research Question 2: Are Liars Less Certain Than Truth-Tellers?	
12	**Tentative**	% of tentative words (e.g., maybe, perhaps, see).
13	**Modal Verbs** // Uncertainty // Discrepancy	% of modal verbs or auxiliary verbs or words expressing uncertainty (e.g., should, would, could).
14	**Certainty**	% of words that express certainty (e.g., always, never).
	Research Question 3a: Do Liars Use More Negations and Negative Emotion Words?	
17	**Negations** // Less Positive Tone // Spontaneous Negations // Negation Connectives	% of words that express negations (e.g., no, never, not).
18+	**Negative Emotions** // Negative Affect // Anger // Anxiety, Fear // Sadness	% of words that express negative emotion / affect (e.g., hate, worthless, enemy) AND anger (e.g., hate, kill, annoyed) AND anxiety (e.g., worried, fearful, nervous) AND sadness (e.g., crying, grief, sad).
18.1	**Negative Emotions (only)** // Negative Affect	% of words that express negative emotion / affect (e.g., hate, worthless, enemy).
18.2	**Anger**	% of words that express anger (e.g., hate, kill, annoyed).
18.3	**Anxiety**	% of words that express anxiety (e.g., worried, fearful, nervous).

18.4	**Sadness**	% of words that express sadness (e.g., crying, grief, sad).
	Research Question 3b: Do Liars Use Less Positive Emotion Words?	
19+	**Positive Emotions and Feelings** // Positive Emotions // Positive Affects // Positive Feelings	% of words that express positive emotion / affect (e.g., happy, pretty, good) AND positive feelings (e.g., joy, love).
19.1	**Positive Emotions (only)**// Positive Affect	% of words that express positive emotion / affect (e.g., happy, pretty, good).
19.2	**Positive Feelings (only)**	% of words that express positive feelings (e.g., joy, love).
	Research Question 3c: Do Liars Express More or Less Unspecified Emotion Words?	
15	**Emotions //** Emotional / Affective Processes // Affect (Ratio) // Positive and Negative Affect	% of words that express any type of emotions / affects (e.g., happy, ugly, bitter).
16	**Pleasantness and Unpleasantness**	% of words that express pleasantness / unpleasantness.
	Research Question 4: Do Liars Distance Themselves More From Events?	
20	**Total Pronouns** // Personal Pronouns	% of all personal (e.g., I, our, they) or total pronouns (e.g., that, somebody, the).
21	**First-Person Singular**	% of first-person singular pronouns (e.g., I, my, me).
22	**First-Person Plural**	% of first-person plural pronouns (e.g., we, us, our).
23	**Total First-Person**	% of first-person singular and first-person plural pronouns (e.g., I, we, me).
24	**Total Second-Person**	% of second-person pronouns (e.g., you, you'll).

25	**Total Third-Person** // Other References // Third-Person Singular // Third-Person Plural	% of third-person pronouns (e.g., she, their, them).
26	**Passive Voice Verbs** // Verbal Nonimmediacy	% of passive voice verbs (e.g., “it was searched for”).
27	**Generalizing Terms** // Leveling terms	% of generalizing terms (e.g., everybody, all, anybody).
47	Past Tense Verb	% of past tense verbs (e.g., went, drove, ate).
48	Present Tense Verb	% of present tense verbs of all words (e.g., walk, run, cry).

Research Question 5: Do Liars Use Fewer (Sensory and Contextual) Details?

28+	**Sensory-Perceptual Processes** // Perceptual Processes/Information // Perceptions and Sense // Sensory ratio // See // Hear // Feel	% of words that express sensory-perceptual processes (e.g. taste, touch, feel) AND visual (e.g., view, saw, seen) AND haptical (e.g., feels, touch) AND aural (e.g., listen, hearing) sensory-perceptual processes.
28.1	**Sensory-Perceptual Processes (only)** // Perceptual Processes // Perceptual Information // Perceptions and Sense // Sensory ratio	% of words that express sensory-perceptual processes (e.g. taste, touch, feel).
28.2	**Seeing**	% of words that express visual sensory-perceptual processes (e.g., view, saw, seen).
28.3	**Feeling**	% of words that express tactile sensory-perceptual processes (e.g., feels, touch).
28.4	**Hearing**	% of words that express aural sensory-perceptual processes (e.g., listen, hearing).
29	**Time** // Temporal ratio // Temporal specificity // Temporal cohesion	% of temporal words (e.g., hour, day, o’clock).

30	**Space** // Spatial Terms // Spatial Ratio // Spatial Specificity // Spatial Cohesion	% of spatial words (e.g., around, over, up).
31	**Temporal-spatial Terms** // Temporal and Spatial Details Total // Spatio-Temporal Information // Space and Time	% of temporal (e.g., hour, day, o'clock) AND spatial words (e.g., around, over, up).
32	**Prepositions**	% of prepositions (e.g., on, to, from).
33	**Numbers**	% of numbers (e.g., first, one, thousand).
34	**Quantifier**	% of quantifier (e.g., all, bit, few, less).
35	**Modifiers** (Adverbs & Adjectives) // Rate of Adjectives and Adverbs (Specificity and Expressiveness)	% of modifier: adverbs & adjectives (e.g., here, much, few, very).
36	**Motion Verbs** // Motion Terms	% of words that describe movements (e.g., walk, move, go).
	Research Question 6: Do liars refer less often to cognitive processes?	
37	**Cognitive Processes** // All Connectives	% of words related to cognitive processes (e.g., cause, know, ought).
38	**Insight**	% of words related to a person's insight (e.g., think, know, consider).

Notes. **Bold** font indicate the name of the linguistic cue chosen for this meta-analysis; * No ratio; + indicates that the specific linguistic cue is an umbrella term; % = number of specific words divided by total number of words.

Appendix B

Definition of Linguistic Cues to Deception--Miscellaneous Category

	Linguistic Cue	Final Operational Definition
39	Redundancy	Ratio of function words to number of sentences. Function words, such as articles and pronouns, are used to form grammatical relationships between other words. // The ratio of the number of function words to the number of messages. // Repetitive words. // Argument overlap: Explicit overlap between two sentences by tracking the common nouns in either single or plural form.
40	Assent	% of words that express an assent (e.g., agree, ok, yes).
41	Articles	% of articles (e.g., a, lot, an, the).
42	Inhibition	% of words that express inhibition (e.g., block, constrain, stop).
43	Social Processes	% of words that express social processes e.g., (talk, us, friend).
44	Friends	% of words that are related to friends (e.g., buddy, friend, neighbor).
45	Family	% of words that are related to family (e.g., daughter, husband, aunt).
46	Humans	% of words that are related to humans (e.g., adult, baby, boy).
49	Future Tense Verb	% of future tense verbs (e.g., will, going to).
50	Inclusive	% of inclusive words (e.g., with, and, include).
51	Achievement	% of words that express achievement (e.g., earn, hero, win).
52	Leisure	% of words that express leisure activities (e.g., cook, chat, movie).

53	Emotiveness	Total number of adjectives and total numbers of adverbs divided by total number of nouns and total numbers of verbs.
54	Pausality	Total number of punctuation marks divided by total number of sentences.
55	Swear Words	% of swear words (e.g., ass, heck, shit).
56	Biology	% of words that express biological processes/states (e.g., eat, pain, wash).
57	Health	% of words that express health issues (e.g., hospital, pill, flu).
58	Sexual	% of words that express sexual activities/states (e.g., passion, rape, sex).
59	Optimism	% of words that express optimism (e.g., certainty, pride, win).
60	Communication	% of words that express communication (e.g., talk, share, converse).
61	Occupation	% of words that express occupation (e.g., work, class, boss).
62	School	% of words that express school issues (e.g., class, student, college).
63	Job / Work	% of words that express job issues (e.g., employ, boss, career).
64	Home	% of words that express home issues (e.g., bed, home, room).
65	Sports	% of words that express sport (e.g., football, game, play).
66	Money	% of words that express money and financial issues (e.g., cash, taxes, income).
67	Physical	% of words that express physical states and functions (e.g., ache, breast, sleep).
68	Body	% of words that express body states and symptoms (e.g., asleep, heart, cough).
69	Eating	% of words that express eating, drinking, dieting issues (e.g., eat, swallow, taste).

Notes. % = number of specific words divided by total number of words.

Appendix C

Coding Decisions for Moderator Variables for Each Study

Authors (Year)	Publ. Type	Program	Lang.	Theory	Select.	Age	Prepa-ration	Event Type	Valence	Inter-action	Moti-vation	Mode
Ali & Levine (2008, denials)	publ.	LIWC01	E	IDT/RM	a-priori	adults	n/a	mock crime	neg.	interview	low	spoken
Ali & Levine (2008, confess.)	publ.	LIWC01	E	IDT/RM	a-priori	adults	n/a	mock crime	neg.	interview	low	spoken
Almela et al. (2012)	publ.	LIWC01	S	none	a-priori	adults	n/a	att./liking	neg./pos.	none	low	typed
Bedwell et al. (2011)	publ.	Coh-Metrix	E	other	sign.	adults	prep.	trivial LE	neutral	instruct.	low	spoken
Bond & Lee (2005)	publ.	LIWC01	E	IDT/RM	a-priori	adults	prep.	video	neg.	interact.	low	spoken
Brunet (2009)*	Thesis	LIWC01	E	LIWC	a-priori	child.	n/a	sign. LE	neg./pos.	interview	none	spoken
Burgoon & Qin (2006)	publ.	GATE	E	IDT/RM	a-priori	adults	n/a	other	neutral	interview	low	spoken
Chen (2010; Dataset 3)	Diss.	LIWC01	E	IDT/RM	a-priori	adults	n/a	n/a	neutral	none	n/a	typed
Colwell et al. (2002)	publ.	Wordscan	E	other	a-priori	adults	n/a	live	neg.	interview	n/a	n/a
Cooper (2008)	Diss.	Connexor+	E	IDT/RM	a-priori	adults	n/a	other	neutral	n/a	n/a	typed
Derrick et al. (2012)	pres.	ADAM	E	IDT/RM	a-priori	adults	n/a	other	neutral	none	low	typed
Dzindolet & Pierce (2005)	pres.	LIWC01	E	LIWC	a-priori	adults	n/a	att./liking	neg./pos.	instruct.	n/a	written

Evans et al. (2012, Interv. 1)	publ.	LIWC01	E	LIWC	a-priori	child.	n/a	other	pos.	interview	none	n/a
Fuller et al. (2006, Agent99A.)	pres.	Agent99A.+	E	IDT/RM	a-priori	adults	n/a	real case	neg.	n/a	high	written
Hancock et al. (2008) / Duran et al. (2010)	publ.	LIWC01/ Coh-Metrix	E	IDT/RM	a-priori	adults	prep.	trivial LE	neg./pos.	cmc	n/a	typed
Humpherys et al. (2011)	publ.	Agent99A.	E	IDT/RM	a-priori	adults	n/a	real case	neutral	none	high	typed
Jensen et al. (2011)	publ.	LIWC01	E	other	a-priori	adults	n/a	real case	neg.	interview	high	n/a
Koyanagi & Blandón-Gitlin (2011)	pres.	LIWC07	E	IDT/RM	a-priori	child.	no	other	neutral	interview	none	spoken
Krachow (2010)	publ.	LIWC07	E	IDT/RM	a-priori	adults	prep.	trivial LE	neg./pos.	instruct.	none	spoken
Lee et al. (2009)	publ.	LIWC01	E	other	a-priori	adults	n/a	other	neutral	cmc	low	typed
Liu et al. (2012)	pres.	LIWC07+	E	LIWC	a-priori	adults	n/a	real case	neg.	interact.	high	n/a
Masip et al. (2012)	publ.	LIWC07	S	IDT/RM	a-priori	adults	no	trivial LE	pos.	instruct.	low	written
Morgan et al. (2011, free recall)	publ.	"automated analysis method"	E	none	a-priori	adults	n/a	n/a	neg.	interview	low	spoken
Morgan et al. (2008, free recall)	publ.	n/a (general)	A	none	a-priori	adults	no	mock crime	neg.	interview	med.	spoken
Newman et al. (2003, Exp. 1)	publ.	LIWC01	E	LIWC	sign.	adults	n/a	att./liking	neutral	instruct.	low	spoken

Newman et al. (2003, Exp. 2)	publ.	LIWC01	E	LIWC	sign.	adults	n/a	att./liking	neutral	instruct.	low	typed
Newman et al. (2003, Exp. 3)	publ.	LIWC01	E	LIWC	sign.	adults	n/a	att./liking	neutral	none	low	written
Newman et al. (2003, Exp. 4)	publ.	LIWC01	E	LIWC	sign.	adults	n/a	att./liking	neg./pos.	instruct.	low	spoken
Newman et al. (2003, Exp. 5)	publ.	LIWC01	E	LIWC	sign.	adults	n/a	mock crime	neg.	interview	low	spoken
Ott et al. (2011)	pres.	LIWC07	E	IDT/RM	sign.	adults	n/a	att./liking	pos.	none	low	typed
Qin et al. (2005, audio)	pres.	GATE	E	other	a-priori	adults	n/a	mock crime	neg.	interview	low	spoken
Qin et al. (2005, face-to-face)	pres.	GATE	E	other	a-priori	adults	n/a	mock crime	neg.	interview	low	n/a
Qin et al. (2005, text chat)	pres.	GATE	E	other	a-priori	adults	n/a	mock crime	neg.	interview	low	typed
Rowe & Blandón-Gitlin (2008)	pres.	LIWC07	E	IDT/RM	a-priori	adults	no	other	neutral	interview	none	spoken
Schafer (2007, Exp. 1)	Diss.	MS Word	E	other	a-priori	adults	n/a	video	neg.	n/a	none	written
Schafer (2007, Exp. 2)	Diss.	MS Word	E	other	a-priori	adults	n/a	video	neg.	n/a	none	written
Schelleman-Offermans & Merckelbach (2010)	publ.	LIWC01	D	LIWC	a-priori	adults	n/a	sign. LE	neg.	n/a	none	typed
Suckle-Nelson et al. (2011, free recall)	publ.	Wordscan	E	IDT/RM	a-priori	adults	n/a	staged	neg.	interview	none	spoken
ten Brinke & Porter (2012)	publ.	LIWC01	E	LIWC	a-priori	adults	n/a	real case	neg.	interact.	high	spoken

Van Swol et al. (2012)	publ.	LIWC07+	E	IDT/RM	a-priori	adults	n/a	other	neutral	interact.	med.	spoken
Williams et al. (2012)	subm.°	LIWC07	E	LIWC	a-priori	child.	n/a	sign. LE	neutral	interview	none	spoken
Zhou (2005)	publ.	"NLP tool"	E	other	a-priori	adults	n/a	other	neutral	cmc	low	typed
Zhou et al. (2004)	publ.	iSkim/CueCal	E	other	a-priori	adults	n/a	other	neutral	cmc	none	typed
Zhou & Zhang (2004)	pres.	"Message analyzing software"	E	other	a-priori	adults	n/a	other	neutral	cmc	low	typed

Notes. confess. = confessions; Publ. = Publication; publ. = published; pres. = presented (Poster or Paper); subm. = submitted; Diss. = Dissertation; LIWC01 = LIWC 2001; LIWC07 = LIWC 2007; GATE = General Architecture for Text Engineering; ADAM = Automated Deception Analysis Machine; Agent99A. = Agent99Analyzer; MS Word = Microsoft Word; NLP = natural language processing; + = Study additionally applied a second program; Lang. = Language; A = Arabic; D = Dutch; E = English; S= Spanish; n/a = not available; IDT = Interpersonal Deception Theory; RM = reality monitoring; Select. = Selection; child. = children; Prepar./prep. = preparation; att. = attitude; staged = staged event; sign. = significant; LE = life events; cmc = computer-mediated communication; instruct. = instruction; med. = medium; * = In the meantime, Brunet, Evans, Talwar, Bala, Lindsay, and Lee (2013) formally published the data of Brunet's Thesis; ° = In the meantime, Williams, Talwar, Lindsay, Bala & Lee (2012) published their (at the time of conducting the meta-analyses unpublished) manuscript.

Appendix D

Meta-Analyses on Miscellaneous Linguistic Cues with and without Outliers

Linguistic Cue	k	N	Min g_u	Max g_u	g_u	CI-low	CI-high	Q	I^2
39 Redundancy	9	1,262	-0.33	0.42	0.00	-0.12,	0.12	9.12	12.30
40 Assent	12	2,452	-0.23	0.38	-0.01	-0.09,	0.07	12.50	12.02
41 Articles	14	2,479	-1.95	0.26	-0.01	-0.08,	0.08	**39.49**	67.08
41 Articles[wO]	12	1,777	-1.95	0.15	-0.02	-0.11,	0.07	18.68	41.12
42 Inhibition	12	2,452	-0.37	0.31	**0.12**	0.04,	0.20	19.27	42.91
43 Social Processes	15	2,979	-1.69	0.26	**-0.26**	-0.33,	-0.18	**236.03**	94.07
43 Social Processes[wO]	14	2,479	-0.48	0.26	-0.04	-0.12,	0.04	16.97	23.39
44 Friends	11	2,344	-0.47	0.51	0.08	-0.01,	0.16	**23.89**	58.15
44 Friends[wO]	9	1,734	-0.47	0.39	0.00	-0.10,	0.09	11.95	33.07
45 Family	10	2,284	-0.39	0.36	-0.03	-0.11,	0.06	**30.77**	70.75
45 Family[wO]	9	1,784	-0.16	0.36	0.07	-0.02,	0.17	10.09	20.70
46 Humans	11	2,344	-0.45	0.42	0.03	-0.06,	0.11	**27.85**	64.09
46 Humans[wO]	9	1,734	-0.27	0.42	**0.12**	0.02,	0.21	12.82	37.58
49 Future Tense	15	2,979	-1.00	0.48	**-0.24**	-0.31,	-0.16	**98.26**	85.75
49 Future Tense[wO]	14	2,497	-0.35	0.48	**-0.10**	-0.18,	-0.02	22.02	40.97
50 Inclusive	15	2,979	-1.00	**0.16**	-0.16	-0.23,	-0.09	**106.73**	86.88
50 Inclusive[wO]	14	2,497	-0.43	0.16	-0.01	-0.01,	0.07	13.77	5.60
51 Achievement	11	2,344	-0.55	0.46	0.04	-0.04,	0.12	17.96	44.31
52 Leisure	11	2,344	-0.41	0.21	-0.05	-0.13,	0.03	**24.37**	58.97

(*Appendix D continues*)

Appendix D (*continued*)

Linguistic Cue	*k*	*N*	Min g_u	Max g_u	g_u	CI-low	CI-high	*Q*	I^2
52 Leisure[wO]	10	1,844	-0.41	0.11	**-0.12**	-0.21,	-0.03	13.96	35.54
53 Emotiveness	9	1,158	-0.39	0.27	0.04	-0.09,	0.16	9.06	22.74
54 Pausality	8	1,158	-0.31	0.75	-0.09	-0.21,	0.04	**16.57**	57.76
54 Pausality[wO]	7	1,128	-0.31	0.46	-0.11	-0.24,	0.01	11.23	46.59
55 Swear Words	10	2,284	-0.17	0.31	-0.04	-0.12,	0.04	5.11	0.00
56 Biology	4	1,198	-0.41	0.40	**0.16**	0.05,	0.28	**9.87**	69.61
57 Health	4	1,198	-0.11	0.22	0.05	-0.06,	0.16	6.51	53.89
58 Sexual	9	2,190	-0.31	0.57	0.08	-0.01,	0.16	12.02	33.42
59 Optimism	8	1,254	-0.28	0.30	0.01	-0.10,	0.12	9.02	22.40
60 Communication	8	1,254	-0.25	0.20	0.07	-0.04,	0.18	2.87	0.00
61 Occupation	7	1,146	-0.39	0.09	-0.07	-0.18,	0.05	4.87	0.00
62 School	7	1,146	-0.16	0.28	0.03	-0.08,	0.15	4.12	0.00
63 Job	11	2,344	-0.31	0.20	0.02	-0.06,	0.11	9.27	0.00
64 Home	11	2,344	-0.36	0.36	0.03	-0.05,	0.11	**24.06**	58.43
64 Home[wO]	10	1,844	-0.36	0.36	-0.03	-0.12,	0.06	16.64	45.91
65 Sports	7	1,146	-0.55	0.08	-0.08	-0.19,	0.04	8.16	26.44
66 Money	12	2,446	-0.85	0.43	-0.01	-0.08,	0.08	**33.62**	67.28
66 Money[wO]	11	2,344	-0.41	0.43	0.03	-0.05,	0.11	17.96	45.33
67 Physical	7	1,146	0.03	0.31	**0.15**	0.04,	0.27	1.52	0.00
68 Body	11	2,344	-0.39	0.44	0.02	-0.07,	0.10	13.40	25.38
69 Eating	7	1,146	-0.08	0.59	**0.12**	0.01,	0.24	7.85	23.54

Notes. Please consult notes from Table 1.

Appendix E

Effect Sizes of Linguistic Cues to Deception when Studies were Unpublished or Published

Linguistic Cue	*k*	Overall g_u [CI]	k_1	Unpublished	k_2	Published
01 Word Quantity	42	0.24 [0.19, 0.29]	17	***0.27 [0.21, 0.34]***	25	**0.19 [0.11, 0.27]**
03 Type-Token Ratio	22	0.14 [0.74, 0.21]	9	***0.24 [0.15, 0.33]***	13	0.03 [-0.07, 0.13]
15 Emotions	25	-0.34 [-0.41, -0.28]	10	***-0.56 [-0.65, -0.48]***	15	**-0.11 [-0.20, -0.02]**
18.1 Negative Emotions	24	-0.18 [-0.24, -0.11]	8	***-0.24 [-0.33, -0.14]***	16	**-0.12 [-0.21, -0.03]**
20 Total Pronouns[wO]	18	0.29 [0.20, 0.37]	7	**0.45 [0.33, 0.58]**	11	**0.13 [0.01, 0.25]**
21 First-Person Singular	21	-0.06 [-0.13, 0.00]	8	**-0.29 [-0.38, -0.20]**	14	***0.19 [0.10, 0.29]***
23 Total First-Person[wO]	22	0.14 [0.06, 0.22]	8	0.05 [-0.08, 0.18]	14	***0.20 [0.10, 0.30]***
24 Total Second-Person	23	-0.16 [-0.23, -0.10]	12	***-0.21 [-.29, -0.13]***	11	-0.09 [-0.19, -0.01]
25 Total Third-Person	29	-0.21 [-0.27, -0.15]	13	***-0.31 [-0.38, -0.23]***	16	-0.08 [-0.17, -0.01]
28.1 Sens.-Percept. Processes[+]	27	0.06 [0.00, 0.13]	11	0.00 [-0.09, 0.09]	16	***0.13 [0.04, 0.22]***
30 Space	23	0.00 [-0.07, 0.06]	10	0.07 [-0.03, 0.16]	13	-0.07 [-0.16, 0.02]
36 Motion Verbs[wO]	16	-0.10 [-0.17, 0.01]	5	0.01 [-0.12, 0.14]	11	**-0.15 [-0.26, -0.05]**

Notes. Please consult notes from Table 2.

Appendix F

Effect Sizes of Linguistic Cues to Deception when Studies Applied either a Between- or Within-Participants Design

Linguistic Cue	*k*	Overall g_u [CI]	k_1	Between-Participants	k_2	Within-Participants
01 Word Quantity	42	0.24 [0.19, 0.29]	24	0.07 [-0.01, 0.15]	18	***0.36 [0.30, 0.43]***
03 Type-Token Ratio	22	0.14 [0.07, 0.21]	16	***0.33 [0.24, 0.42]***	6	-0.08 [-0.18, 0.02]
12 Tentative Words[wO+]	19	0.13 [0.06, 0.20]	9	0.03 [-0.09, 0.14]	10	***0.19 [0.11, 0.28]***
15 Emotions	25	-0.35 [-0.41, -0.28]	14	**-0.50 [-0.58, -0.40]**	11	**-0.22 [-0.30, -0.13]**
17 Negations	20	-0.15 [-0.22, -0.08]	7	**0.17 [0.03, 0.30]**	13	***-0.25 [-0.32, -0.17]***
18 Negative Emotions[+]	24	-0.13 [-0.19, -0.06]	11	0.04 [-0.08, 0.15]	13	***-0.22 [-0.29, 0.13]***
18.1 Negative Emotions Only	24	-0.18 [-0.24, -0.11]	11	0.09 [-0.02, 0.20]	13	***-0.32 [-0.40, -0.24]***
19 Positive Emotions and Feelings[+]	20	-0.04 [-0.13, 0.04]	11	**-0.16 [-0.27, -0.05]**	9	0.13 [-0.01, 0.26]
19.1 Positive Emotions Only[wO]	20	-0.07 [-0.15, 0.00]	11	**-0.15 [-0.26, -0.04]**	9	0.00 [-0.11, 0.10]
20 Total Pronouns[wO]	18	0.29 [0.10, 0.38]	8	-0.05 [-0.23, 0.14]	10	**0.38 [0.28, 0.48]**
21 First-Person Singular	22	-0.06 [-0.13, 0.00]	9	**-0.70 [-0.80, -0.59]**	13	***0.28 [0.20, 0.36]***
22 First-Person Plural	25	-0.04 [-0.10, 0.03]	15	**-0.16 [-0.25, -0.08]**	10	***0.09 [0.01, 0.18]***
23 Total First-Person[wO]	22	0.14 [0.06, 0.22]	14	0.00 [-0.12, 0.12]	8	***0.26 [0.15, 0.37]***
24 Total Second-Person	23	-0.16 [-0.23, -0.10]	**10**	***-0.42 [-0.52, -0.32]***	13	0.00 [-0.08, 0.08]
25 Total Third-Person	29	-0.21 [-0.27, -0.15]	**15**	***-0.34 [-0.42, -0.25]***	14	**-0.11 [-0.19, -0.03]**
48 Present Tense[wO]	16	0.01 [-0.06, 0.09]	7	-0.10 [-0.24, 0.04]	9	0.06 [-0.03, 0.15]
28.1 Sens.-Percept. Processes Only	27	0.06 [0.00, 0.13]	15	-0.04 [-0.13, 0.05]	12	***0.14 [0.06, 0.23]***
30 Space	24	0.00 [-0.06, 0.07]	**12**	***0.21 [0.11, 0.32]***	12	**-0.12 [-0.20, -0.04]**
37 Cognitive Processes[wO]	18	0.09 [0.01, 0.16]	7	-0.03 [-0.17, 0.10]	11	***0.13 [0.05, 0.22]***
38 Insight	15	0.13 [0.05, 0.21]	7	-0.07 [-0.21, 0.06]	8	***0.23 [0.14, 0.32]***

Notes. Please consult notes from Table 2.

Appendix G

Excluded Studies and Reason for Exclusion

Authors	Reason for Exclusion
Adams (2002)	No additional data from authors
Adams & Jarvis (2006)	No additional data from authors
Bachenko, Fitzpatrick & Schonwetter (2008)	Data not applicable
Burgoon, Blair, Qin, & Nunamaker (2003)	Same data as Qin, Burgoon, Blair, & Nunamaker (2005)
Churyk, Lee, & Clinton (2008)	Truth status of management discussion not clear (could be fraud), no data
Dilmon (2009)	Response from author: No computer program used. A statistician conducted the coding.
Dulaney (1982)	Data (means, *F*-value) are not sufficient to calculate appropriate ES for within-participants design
Duran, Crossley, Hall, McCarthy, & Namara (2009)	Same data as Duran, Hall, McCarthy, & McNamara (2010)
Dzindolet & Pierce (2004)	First author could not provide data
Elkins (2011)	Data not applicable
Enos (2009)	Not enough data to calculate effect sizes
Enos, Shribergy, Graciarena, Hirschberg, & Stolcke (2007)	Not enough data to calculate effect sizes
Fornaciari & Poesio (2011)	Parts of speech instead of whole account. Not enough statistical data (use of "vectors")
Fuller (2008)	Data not applicable
Fuller, Biros, & Delen (2008)	Data not applicable
Fuller, Biros, & Delen (2011)	Not enough data to calculate effect sizes, no independent dataset (same as Fuller, Biros, & Wilson, 2009)

(*Appendix G continues*)

Appendix G (*continued*)

Authors	Reason for Exclusion
Fuller, Biros, & Wilson (2009)	Not enough data and no independent dataset
Graciarena, Shriberg, Stolcke, Enos, Hirschberg, & Kajarekar (2006)	No linguistic categories outlined specifically - used a superordinate category (Prosodic/Lexical)
Gupta (2007)	Not enough data to calculate effect sizes
Hancock, Curry, Goorha, & Woodworth (2004)	Exactly the same data as in Hancock et al. (2008)
Hancock, Curry, Goorha, & Woodworth (2005)	Exactly the same data as in Hancock et al. (2008)
Hirschberg, Beus, Brenier, Enos, Friedman, Gilman, Girand, Graciarena, Kathol, Michaelis, Pellom, Shriberg, & Stolcke (2005)	Not enough data to calculate effect sizes
Jensen, Burgoon, & Nunamaker (2010)	Hybrid detection system whereby humans interact with the program to aid them in making truth and deception decisions
Jensen, Lowry, Burgoon, & Nunamaker (2010)	Hybrid detection system whereby humans interact with the program to aid them in making truth and deception decisions
Jensen, Lowry, & Jenkins (2011)	Hybrid detection system whereby humans interact with the program to aid them in making truth and deception decisions
Jensen, Meservy, Burgoon, & Nunamaker (2010)	No independent data. Same transcripts as in Burgoon, et al. (2003)
Keila & Skillicorn (2005)	Data not applicable
Knapp, Hart, & Dennis (1974)	No standard deviations reported and first author could not provide them
Leuprecht (2011)	No data available
Little (2007)	Not enough data to calculate effect sizes

(*Appendix G continues*)

Appendix G (*continued*)

Authors	Reason for Exclusion
Mihalcea & Strappavara (2009)	Not enough data and no independent dataset
Morgan, Rabinowitz, Christian, & Hazlett (2009)	Analyses of interviewers speech only.
Morgan, Steffian, Clark, Coric, & Harzlett (2008)	No data available
Qin & Burgoon (2007)	Same dataset as Burgoon & Qin (2006)
Qin, Burgoon, & Nunamaker (2004)	Same data as Qin, Burgoon, Blair, & Nunamaker (2005)
Rubin & Conroy (2012)	Analysis of statements with various (continuous and self-determined) deception levels and no restricted topics. No data available
Taylor, Tomblin, Conchie, & Menacere (2011)	Not enough data to calculate effect sizes
Toma & Hancock (2010)	No truth condition, only low and high deceptive condition, and no data. Same dataset as Toma & Hancock (2012)
Toma & Hancock (2012)	No truth condition, only low and high deceptive condition, and no data
Twitchell (2005)	Same data as Twitchell, Adkins, Nunamaker, & Burgoon (2004)
Twitchell, Nunamaker, & Burgoon (2004)	Only one linguistic cue investigated
Twitchell, Biros, Adkins, Forsgren, Burgoon, Nunamaker (2006)	Data not applicable
Vrij, Mann, Kristen, & Fisher (2007)	First author could not provide data
Watson (1981)	Statistical data is not useful for computing effect sizes
Zhou, Burgoon, Twitchell, Qin, & Nunamaker (2004)	No additional data from authors
Zhou, Burgoon, Zhang, & Nunamaker (2004)	Same data as Zhou et al. (2004), but two additional dependent variables: "intensity" and "subjunctive language". Data for those cues is neither provided in the article, nor in data sent from the authors.

(*Appendix G continues*)

Appendix G (*continued*)

Authors	Reason for Exclusion
Zhou, Shi, & Zhang (2008)	Data not applicable
Zhou & Sung (2008)	First author recommend us to exclude this study because it does not have an independent data set
Zhou, Twitchell, Qin, Burgoon, & Nunamaker (2003)	First author recommend us to exclude this study because it does not have an independent data set
Zhou & Zenebe (2005)	Data not applicable
Zhou & Zenebe (2008)	Data not applicable
Zhou & Zhang (2006)	First author recommend us to exclude this study because it does not have an independent data set

Appendix H

Meta-Analyses of Linguistic Cues under Research Questions 1 to 6 with and without Outliers

Linguistic Cue	Pred. DOE	*k*	*N*	Min g_u	Max g_u	g_u	CI-low	CI-high	*Q*	I^2
Research Question 1: Do Liars Experience Greater Cognitive Load?										
(a) Length of Accounts										
01 Word Quantity[M]	T	42	6,713	-1.25	1.43	**0.24**	0.19,	0.29	**315.85**	87.02
07 Sentence Quantity	T	9	1,334	-1.31	0.28	**-0.33**	-0.44,	-0.21	**104.01**	93.31
08 Average Sentence Length	T	16	2,880	-0.37	0.81	**0.10**	0.02,	0.17	**42.83**	64.98
08 Average Sentence Length[wO,M]	T	15	2,704	-0.37	0.43	0.05	-0.03,	0.13	20.46	31.58
(b) Elaboration of Accounts										
02 Content Word Diversity	T	9	1,194	-0.30	1.00	**0.39**	0.26,	0.51	**22.77**	64.87
02 Content Word Diversity[wO]	T	7	1,076	0.27	1.00	**0.48**	0.34,	0.61	8.13	26.22
03 Type-Token Ratio[M]	T	22	3,589	-1.40	1.09	**0.14**	0.07,	0.21	**171.95**	87.79
04 Six Letter Words	T	10	1,617	-0.28	0.25	-0.05	0.14,	-0.05	5.19	0.00
05 Average Word Length	T	8	1,158	-0.59	0.69	-0.03	-0.16,	0.09	**25.95**	73.02
05 Average Word Length[wO]	T	7	954	-0.42	0.69	0.11	-0.03,	0.25	6.85	12.41
(c) Complexity of Accounts										
06 Verb Quantity	T	12	2,356	-1.21	0.44	-0.03	0.11,	-0.60	**78.91**	86.06
09 Causation	T	17	2,773	-0.68	0.25	-0.07	-0.14,	0.01	18.61	14.03
10 Exclusive Words	T	20	3,403	-0.24	0.81	**0.31**	0.24,	0.38	**43.91**	56.73
10 Exclusive Words[wO,M]	T	18	2,783	-0.24	0.81	**0.24**	0.17,	0.31	25.87	25.66

(*Appendix H continues*)

Appendix H (*continued*)

Linguistic Cue	Pred. DOE	*k*	*N*	Min g_u	Max g_u	g_u	CI-low	CI-high	*Q*	I^2
(d) Errors in Production										
11 Writing Errors	D	10	1,077	-0.65	0.86	-0.03	-0.16,	0.11	**22.10**	59.29
11 Writing Errors^{wO}	D	8	990	-0.65	0.43	-0.01	-0.15,	0.12	13.36	47.61
Research Question 2: Are Liars Less Certain Than Truth-Tellers?										
12 Tentative Words	D	20	3,197	-1.27	0.36	**0.11**	0.04,	0.18	**36.76**	48.31
12 Tentative Words^{wO}	D	19	3,145	-1.27	0.36	**0.13**	0.06,	0.20	20.13	10.56
13 Modal Verbs	D	25	3,889	-0.42	0.80	0.00	-0.07,	0.06	32.73	26.62
14 Certainty	T	18	2,823	-0.25	0.94	-0.06	-0.14,	0.01	25.15	32.40
Research Question 3a: Do Liars Use More Negations and Negative Emotion Words?										
17 Negations^{M}	D	20	3,659	-0.98	0.53	**-0.15**	-0.22,	-0.09	**155.53**	87.78
18 Negative Emotions^{+}	D	24	3,641	-1.90	0.87	**-0.13**	-0.19,	-0.06	**99.80**	76.95
18 Negative Emotions$^{+,wO,M}$	D	21	2,593	-0.39	0.87	-0.07	-0.15,	0.01	21.03	4.88
18.1 Negative Emotions Only^{M}	D	24	3,641	-1.90	0.87	**-0.18**	-0.24,	-0.12	**214.57**	89.28
18.2 Anger	D	12	2,452	-1.32	0.38	**-0.27**	-0.35,	-0.19	**165.03**	93.34
18.3 Anxiety^{wO}	D	11/12	1,952	-0.30	0.44	0.07	-0.02,	0.02	13.55	26.19
18.4 Sadness	D	12	2,452	-0.34	0.25	0.04	-0.04,	0.12	13.01	15.46
Research Question 3b: Do Liars Use Fewer Positive Emotion Words?										
19 Positive Emotions and Feelings^{+}	T	21	3,203	-0.84	0.45	0.03	-0.04,	0.10	**55.42**	63.91
19 Positive Emotions and Feelings$^{+,wO}$	T	20	2,703	-0.84	0.37	-0.05	-0.12,	0.03	29.59	35.79

(*Appendix H continues*)

Appendix H (*continued*)

Linguistic Cue	Pred. DOE	*k*	*N*	Min g_u	Max g_u	g_u	CI-low	CI-high	*Q*	I^2
19.1 Positive Emotions Only	T	21	3,203	-0.84	0.45	0.01	-0.06,	0.08	**56.68**	64.71
19.1 Positive Emotions Only[wO,M]	T	20	2,703	-0.84	0.35	-0.07	-0.15,	0.00	27.98	32.08
19.2 Positive Feelings Only	T	9	1,422	-0.47	0.40	0.07	-0.03,	0.18	14.88	46.25
Research Question 3c: Do Liars Express More or Less Unspecified Emotion Words?										
15 Emotions	?	25	4,129	-1.57	0.48	**-0.34**	-0.41,	-0.28	**246.69**	90.27
15 Emotions[wO,M]	?	21	2,941	-0.63	0.48	**-0.11**	-0.19,	-0.04	28.92	30.85
16 Pleasantness and Unpleasantness	?	6	806	-0.35	0.30	-0.10	-0.25,	0.06	8.43	40.68
Research Question 4: Do Liars Distance Themselves More From Events?										
(a) Personal Pronouns										
21 First-Person Singular[M]	T	22	3,761	-1.00	0.61	-0.06	-0.13,	0.00	**274.38**	92.35
22 First-Person Plural	T	25	4,353	-1.12	0.79	-0.04	-0.10,	0.02	**228.85**	89.51
22 First-Person Plural[wO,M]	T	22	3,224	-0.72	0.39	0.06	-0.01,	0.13	27.98	24.93
23 Total First-Person	T	23	2,709	-1.63	0.57	0.05	-0.03,	0.13	**127.30**	82.72
23 Total First-Person[wO,M]	T	22	2,541	-0.39	0.57	**0.14**	0.06,	0.22	32.26	34.91
24 Total Second-Person	D	23	4,072	-1.18	0.33	**-0.16**	-0.23,	-0.10	**168.24**	86.92
24 Total Second-Person[wO,M]	D	21	3,072	-0.61	0.33	**-0.10**	-0.17,	-0.02	28.64	30.16
25 Total Third-Person	D	29	4,807	-1.18	0.55	**-0.21**	-0.27,	-0.15	**181.71**	84.59
25 Total Third-Person[wO,M]	D	26	3,848	-0.41	0.55	**-0.10**	-0.16,	-0.04	37.20	32.79
20 Total Pronouns	-	19	2,960	-0.64	0.65	-0.06	-0.13,	0.02	**68.18**	73.60
20 Total Pronouns[wO,M]	-	18	2,460	-0.36	0.65	0.06	-0.02,	0.14	19.32	12.02

(*Appendix H continues*)

Appendix H (*continued*)

Linguistic Cue	Pred. DOE	k	N	Min g_u	Max g_u	g_u	CI-low	CI-high	Q	I^2
(b) Passive Voice and Generalizing Terms										
26 Passive Voice Verbs	D	11	1,221	-0.47	0.49	0.06	-0.06,	0.18	9.52	0.00
27 Generalizing Terms	D	4	93	-1.63	0.44	-0.37	-0.79,	0.05	**15.78**	80.99
(c) Past and Present Tense										
47 Past Tense	D	16	3,047	-0.53	0.41	0.06	-0.01,	0.14	22.67	33.83
48 Present Tense	T	17	3,107	-1.41	0.60	**-0.18**	-0.25,	-0.11	**195.37**	91.81
48 Present Tense^{wO,M}	T	16	2,607	-0.51	0.60	0.01	-0.07,	0.09	19.38	22.61
Research Question 5: Do Liars Use Fewer (Sensory and Contextual) Details?										
28 Sens.-Perceptual Processes^{+}	T	27	4,177	-0.70	0.70	**0.06**	0.00,	0.13	**58.84**	55.81
28 Sens.-Perceptual Processes$^{+,wO,M}$	T	25	3,957	-0.70	0.70	0.05	-0.01,	0.12	36.00	33.33
28.1 Sens.-Percept. Processes Only^{M}	T	27	4,177	-0.90	0.70	**0.06**	0.00,	0.13	**89.26**	70.87
28.2 Seeing	T	11	2,344	-0.56	0.35	0.07	-0.01,	0.16	**33.25**	69.93
28.2 Seeing^{wO}	T	9	1,740	-0.17	0.34	0.03	-0.06,	0.13	13.34	40.03
28.3 Feeling	T	12	2,412	-0.49	0.40	-0.01	-0.09,	0.07	**19.64**	44.00
28.3 Feeling^{wO}	T	11	2,304	-0.49	0.27	-0.03	-0.11,	0.05	15.00	33.31
28.4 Hearing	T	11	2,344	-0.41	0.48	**0.17**	0.09,	0.25	14.15	29.35

(*Appendix H continues*)

Appendix H (*continued*)

Linguistic Cue	Pred. DOE	*k*	*N*	Min g_u	Max g_u	g_u	CI-low	CI-high	*Q*	I^2
(b) Contextual Embedding										
29 Time	T	24	3,796	-1.25	0.53	**0.10**	0.03,	0.16	**53.32**	56.86
29 Time[wO]	T	23	3,296	-1.25	0.53	0.03	-0.04,	0.10	28.95	24.00
30 Space	T	24	3,851	-0.47	0.58	0.00	-0.06,	0.07	**60.12**	61.74
30 Space[wO,M]	T	22	3,199	-0.36	0.58	-0.04	-0.13,	0.03	31.74	33.74
31 Space & Time	T	5	634	-0.25	0.61	-0.04	-0.19,	0.12	**10.48**	61.84
(c) Descriptive Words										
32 Prepositions	T	14	2,479	-0.55	0.48	0.02	-0.06,	0.10	16.54	21.38
33 Numbers	T	12	2,452	-0.28	0.23	0.05	-0.03,	0.13	9.37	0.00
34 Quantifier	T	4	1,198	0.06	0.22	**0.14**	0.02,	0.25	1.22	0.00
35 Modifier	T	11	1,361	-1.04	0.43	-0.08	-0.20,	0.03	**77.46**	87.09
36 Motion Verbs	T	17	2,859	-0.72	0.38	-0.01	-0.08,	0.07	**39.59**	59.59
36 Motion Verbs[wO,M]	T	16	2,359	-0.72	0.13	**-0.09**	-0.17,	-0.01	16.84	10.92
Research Question 6: Do Liars Refer Less Often to Cognitive Processes?										
37 Cognitive Processes	T	19	2,995	-0.25	0.82	**0.10**	0.03,	0.18	**28.97**	37.88
37 Cognitive Processes[wO]	T	18	2,915	-0.25	0.36	**0.09**	0.01,	0.16	19.66	13.54
38 Insight[M]	T	15	2,539	-0.41	0.59	**0.13**	0.05,	0.21	**35.65**	60.73

Notes. Please consult notes from Table 1.