Lying takes time

A meta-analysis on reaction time measures of deception

Suchotzki, K.; Verschuere, B.; Van Bockstaele, B.; Ben-Shakhar, G.; Crombez, G.

DOI
10.1037/bul0000087

Publication date
2017

Document Version
Final published version

Published in
Psychological Bulletin

License
Article 25fa Dutch Copyright Act

Psychological Bulletin
2017, Vol. 143, No. 4, 428–453

0033-2909/17/$12.00 http://dx.doi.org/10.1037/bul0000087

Lying Takes Time: A Meta-Analysis on Reaction Time Measures of Deception

Kristina Suchotzki
University of Würzburg and Ghent University

Bruno Verschuere
University of Amsterdam

Bram Van Bockstaele
Ghent University and University of Amsterdam

Gershon Ben-Shakhar
The Hebrew University of Jerusalem

Geert Crombez
Ghent University

Lie detection techniques are frequently used, but most of them have been criticized for the lack of empirical support for their predictive validity and presumed underlying mechanisms. This situation has led to increased efforts to unravel the cognitive mechanisms underlying deception and to develop a comprehensive theory of deception. A cognitive approach to deception has reinvigorated interest in reaction time (RT) measures to differentiate lies from truths and to investigate whether lying is more cognitively demanding than truth telling. Here, we provide the results of a meta-analysis of 114 studies (n = 3307) using computerized RT paradigms to assess the cognitive cost of lying. Results revealed a large standardized RT difference, even after correction for publication bias (d = 1.049; 95% CI [0.930; 1.169]), with a large heterogeneity amongst effect sizes. Moderator analyses revealed that the RT deception effect was smaller, yet still large, in studies in which participants received instructions to avoid detection. The autobiographical Implicit Association Test produced smaller effects than the Concealed Information Test, the Sheffield Lie Test, and the Differentiation of Deception paradigm. An additional meta-analysis (17 studies, n = 348) showed that, like other deception measures, RT deception measures are susceptible to countermeasures. Whereas our meta-analysis corroborates current cognitive approaches to deception, the observed heterogeneity calls for further research on the boundary conditions of the cognitive cost of deception. RT-based measures of deception may have potential in applied settings, but countermeasures remain an important challenge.

Keywords: deception, lie detection, lying, meta-analysis, reaction time

In February 2012, Trayvon Martin, a 17-year-old African American high school student, was shot. His death sparked a heated debate in the media as well as in criminal court, and President Obama declared: "Trayvon Martin could have been me 35 years ago." George Zimmerman, the neighborhood watch coordinator of the community, admitted to the shooting yet claimed to have acted in self-defense. To determine the veracity of his statement, Zimmerman was taken into custody and interrogated. The interrogation included a lie detection test. Although rarely accepted as evidence in court, lie detection tests are routinely applied by the police in North America (http://www.cvsa1.com/agencies.htm; see also National Research Council, 2003).

The key question in the Zimmerman case was whether he acted in self-defense. The lie test therefore focused on voice stress parameters to two questions: "Did you confront the guy you shot?" and "Were you in fear for your life when you shot the guy?". The voice stress responses to these relevant questions were compared with comparison questions such as "Have you ever driven over the posted speed limit?" (questions retrieved from the video of the original interrogation; https://www.youtube.com/watch?v=S6cN-mTZTjA). In such a lie

This article was published Online First February 9, 2017.

Kristina Suchotzki, Department of Psychology, University of Würzburg, and Department of Experimental-Clinical and Health Psychology, Ghent University; Bruno Verschuere, Department of Clinical Psychology, University of Amsterdam; Bram Van Bockstaele, Department of Experimental-Clinical and Health Psychology, Ghent University, and Departments of Developmental Psychology and Child Development and Education, University of Amsterdam; Gershon Ben-Shakhar, Department of Psychology, The Hebrew University of Jerusalem; Geert Crombez, Department of Experimental-Clinical and Health Psychology, Ghent University.

This research was supported by an ECRP Grant (09-ECRP-025; FWO Grant ESF 3G099310). Kristina Suchotzki worked as a doctoral researcher at Ghent University and is currently working as postdoctoral researcher at the University of Würzburg. Bram Van Bockstaele worked as a postdoctoral researcher at Ghent University and is currently working as postdoctoral researcher of the Research Priority Area YIELD of the University of Amsterdam. The authors thank Jemma Todd, who proofread an earlier version of this manuscript, and all researchers who contributed data and/or information to this meta-analysis.

Correspondence concerning this article should be addressed to Kristina Suchotzki, Department of Psychology, University of Würzburg, Marcusstraße 9-11, 97070 Würzburg, Germany. E-mail: kristina.suchotzki@uni-wuerzburg.de

test, stronger responding to the relevant question compared with the comparison questions is interpreted as an indication of deception. The opposite pattern is considered as an indication of truth telling. Results of the lie test in the Zimmerman case were taken as an indication that he spoke the truth. The police argued to have no reason to doubt the veracity of the self-defense statement, and Zimmerman was released (for the aftermath of the case see, e.g., https://en.wikipedia.org/wiki/Shooting_of_Trayvon_Martin). This case illustrates the enormous potential of lie detectors. If accurate, they can lead to a huge saving of time and money and solve difficult cases quickly. At the same time, the Zimmerman case illustrates three key questions relevant for the use of lie detection methods: (a) How accurate are lie tests? (b) What constitutes a proper baseline to which responses on critical questions can be compared? (c) Which psychological mechanisms underlie deception and its detection?

Accuracies reported for lie tests vary greatly. In 2003, the National Research Council evaluated the validity of the widely used Control Question polygraph test (Reid, 1947) and concluded that

> what is remarkable, given the large body of relevant research, is that claims about the accuracy of the polygraph made today parallel those made throughout the history of the polygraph: Practitioners have always claimed extremely high levels of accuracy, and these claims have rarely been reflected in empirical research. (p. 107)

Importantly, although some lie tests show a predictive validity above chance level, empirical evidence has shown that they perform considerably below perfection (see, e.g., Meijer, Verschuere, Gamer, Merckelbach, & Ben-Shakhar, 2016; National Research Council, 2003). Voice stress analysis as employed in the Zimmermann case does not seem to perform above chance level in discerning lies from truth (Eriksson & Lacerda, 2007).

There is a consensus that there exists no unique lie response (cf. Pinocchio's nose), and therefore evaluating whether a response is truthful or not always requires its comparison to an appropriate "baseline" (Meijer et al., 2016). What constitutes an adequate baseline is, however, heavily debated. In the Zimmerman case, as in most commonly used lie tests, the critical questions and the comparison questions are readily discernable upon face value for both liars and truth tellers. This is undesirable because the critical questions are threatening for all, and may thus elicit a false positive response also in innocent suspects.

Attempts to formulate a psychological theory explaining the expected response pattern in lie detection tests mostly build on the assumption of a unique relation between lying and stress. The basic assumption is that lying leads to stress and enhanced autonomic arousal, and that truth telling will not (or at least less; Ben-Shakhar, 2002; Vrij, 2008). The idea of interpreting stress as a sign of deceit is popular and has a long tradition. Already in 1000 BCE, the observation that fear leads to decreased saliva production was used to identify liars in China. Suspects were asked to chew dried rice and to spit it out. The more difficult this proved to be, the greater the chance that they were considered to be deceptive (Kleinmuntz & Szucko, 1984). Over centuries, the technologies to measure stress have advanced, but investigations of the relation between stress and deception have not. Although it is an intuitively appealing idea that liars display more stress, the possibility of a so-called Othello error (Ekman, 1985) is large: Interpreting his wife Desdemona's stress as a sign of deceit, Shakespeare's Othello failed to consider that stress may also reflect a truth teller's fear of not being believed. In Shakespeare's play, this error in judgment led to the death of Desdemona, illustrating the potentially dramatic consequences for innocent suspects.

Taken together, many currently used lie detection tests perform below perfection, do not use a proper baseline, and are based on the problematic assumption of a strong relationship between stress and deception. Following Kurt Lewin's (1951) suggestion that "There is nothing so practical as a good theory," a very first step in the development of lie detection tools should be the development of a valid theory explaining why and how lying and truth telling differ. Such a theory is necessary, as it would allow to predict under which conditions lie tests are likely to result in valid conclusions and under which conditions they are not. A theoretical perspective that has gained increased attention in recent years is the cognitive view on lying.

The Cognitive Approach to Deception

The cognitive approach to deception holds that lying is cognitively more challenging than truth telling. According to this view, what may differentiate lying from truth telling is the extent and nature of the required cognitive processes. There are several reasons why lying may be cognitively more demanding than truth telling (Vrij et al., 2006; Christ et al., 2009). In comparison with truth tellers, liars may not simply retrieve a story from memory, but need to fabricate it. Liars have to be cautious not to contradict themselves or the knowledge of the person they are trying to deceive. To avoid such contradictions, they have to remember the truth, keep the truth active in memory, and yet at the same time prevent it from slipping out while communicating the lie. Finally, liars have to monitor their own behavior and that of their interaction partners to control behaviors that may be interpreted as lying. In sum, the cognitive view holds that deception is *typically* more cognitively demanding than truth telling (Ellwanger, Rosenfeld, Sweet, & Bhatt, 1996; Johnson, Barnhardt, & Zhu, 2004; Spence et al., 2004; Vrij, 2008; Zuckerman, DePaulo, & Rosenthal, 1981; for an opposing view see, e.g., McCornack, Morrison, Paik, Wisner, & Zhu, 2014; for boundary conditions see, e.g., Walczyk, Harris, Duck, & Mulay, 2014).

There is empirical support for the idea that lying has a cognitive cost. First, people report to experience lying as cognitively more demanding than truth telling (Caso, Gnisci, Vrij, & Mann, 2005; Vrij, Semin, & Bull, 1996). Second, brain imaging research shows that compared with truth telling, lying evokes greater activity in brain regions that are also active during complex cognitive tasks (i.e., dorsolateral prefrontal and inferior frontal regions). In contrast, no brain area has been found to be systematically more active for truth telling compared with lying (Abe, 2009; Christ, Van Essen, Watson, Brubaker, & McDermott, 2009; Gamer, 2011; Ganis & Keenan, 2009).[1] Third, increasing cognitive load in-

[1] The studies by Langleben and colleagues (Davatzikos et al., 2005; Langleben et al., 2005) provide an exception in revealing that some brain areas were more active during truth telling than during lying. However, as will be explained later on in the manuscript, an interpretation of the results reported in these studies is problematic because they relied on a paradigm that confounded the crucial comparison of lying versus truth telling with the ratio of yes versus no responses. Yes responses were infrequent (1:10), and only used for truth trials, whereas No responses were frequent (9:10) and used for lie and filler trials.

creases the occurrence of verbal, nonverbal, and para-verbal (e.g., voice tone, pitch, or pacing) cues of deception. For instance, letting both liars and truth tellers tell their stories in reverse order or asking them unanticipated questions burdens liars in particular, and increases the accuracy of observers in detecting liars (e.g., Lancaster, Vrij, Hope, & Waller, 2013; Vrij et al., 2008, 2009; Warmelink, Vrij, Mann, Jundi, & Granhag, 2012; for reviews of interview-techniques of cognition-based lie-detection, see Vrij, Fisher, & Blank, 2015; Vrij & Granhag, 2012).

Three executive functions (Miyake et al., 2000) have been proposed to underlie the enhanced cognitive cost during deception (Christ et al., 2009; Spence et al., 2001). (a) The formulation of a credible lie often requires the truth to be retrieved and kept active in *working memory*. (b) The to-be-emitted lie contrasts with the truth, and may require *response inhibition* to prevent the truth from coming out and enable a deceptive response instead. (c) The transition from truthful communication to deception (and back) may require *task switching*. Although many studies have looked at these executive functions in isolation, the field is now slowly moving forward to formulate more integrative accounts.

So far, there are a few comprehensive cognitive theories of deception (e.g., Gombos, 2006; McCornack et al., 2014; Spence et al., 2004; Sporer & Schwandt, 2007; Sporer, 2016; Walczyk et al., 2014). The most elaborate theory is the Activation-Decision-Construction-Action Theory (ADCAT) of deception (Walczyk et al., 2014). In this theory, lying involves four components or stages: (a) The activation of the truth from long-term memory, (b) the decision whether and how to deceptively alter the to-be-shared information, (c) the construction of the lie, and (d) the action itself. Importantly, each of these four components may add to the cognitive load during lying. The strength of the ADCAT is that it proposes moderators at each of the four stages, leading to testable hypotheses on the boundary conditions of the cognitive cost of deception.

A key moderator in the ADCAT is motivation. Interpreted as the amount of cognitive resources a liar is willing to recruit, motivation is hypothesized to not only be influential at the decision stage, but also at the construction and action stages, where it may determine how much effort a liar is going to assign to lying successfully. Through its influence on motivation, the relevance of the information a person lies about is also thought to moderate the effort invested in successful lying. Furthermore, Walczyk et al. (2014) proposed that lying becomes more difficult the more entrenched the truthful response is in the liar, whereas it becomes easier the more the deceptive response has been rehearsed and practiced (see also DePaulo et al., 2003). It is theorized that there may in fact be situations in which truth telling turns out to be cognitively more demanding than lying (McCornack et al., 2014; Walczyk et al., 2014). Identifying those situations is crucial to gain information about the validity and the boundaries of cognition-based methods of deception detection.

Measuring Deception With Reaction Times

Departing from the logic of mental chronometry (Donders, 1969; for a review see Jensen, 2006), subtracting reaction times (RTs) for truthful responses from RTs for deceptive responses may provide useful information on the extent to which additional cognitive processes are required for successful lying. Although RTs have a long-standing tradition in experimental psychology, their use in deception research has been contested. Early attempts to employ RTs as deception measures made use of Jung's Association Reaction Method (Jung, 1910). For instance, the Russian researcher Alexander Luria measured the time it took crime suspects to generate word associations to crime-related words compared with crime-unrelated words (Luria, 1932). Interestingly, Luria observed longer RTs and larger RT variability in guilty suspects, which he interpreted as indicators of disturbed affect. He also proposed that his method could differentiate between guilty and innocent suspects. Other researchers have tested variants of the Association Reaction Method, but have failed to observe robust RT differences between lying and truth telling (Henke & Eddy, 1909; Marston, 1920). After these disappointing results, a period of low interest in RTs as deception indices followed (ca. 1940–1990).

Two developments led to renewed interest. First, with the increased use of computerized measures, the measurement of RTs has become easy, accurate, and common. Second, new deception paradigms were developed, grounded in the cognitive approach to deception. Those paradigms aimed at overcoming shortcomings of earlier lie detection tests, for instance by more closely matching the critical (i.e., lie) and the control (i.e., truth) condition (e.g., Furedy, Davis, & Gurevich, 1988; Spence et al., 2001; see method section). Nevertheless, the results of RT studies remained mixed, with effects ranging from small or nonsignificant (e.g., Engelhard, Merckelbach, & van den Hout, 2003; Locker & Pratarelli, 1997; Matsuda, Nittono, Hirota, Ogawa, & Takasawa, 2009; Meijer, Smulders, Merckelbach, & Wolf, 2007; Verschuere, Rosenfeld, Winograd, Labkovsky, & Wiersema, 2009) to very large (e.g., Debey, Verschuere, & Crombez, 2012; Seymour, Seifert, Shafto, & Mosmann, 2000; Seymour & Kerlin, 2008; Spence et al., 2001, 2008; Van Bockstaele et al., 2012; Verschuere, Spruyt, Meijer, & Otgaar, 2011). The practical utility of RTs as deception indices has also been contested. On the one hand, it has been argued that RTs have many advantages over autonomic and neural measures, as they are cheap, easy, and quick to measure (Verschuere & De Houwer, 2011). On the other hand, critics have stated that RTs are under voluntary control and are thus not suited in deception contexts (e.g., Farwell & Donchin, 1991; Gronau, Ben-Shakhar, & Cohen, 2005; Mertens & Allen, 2008; Sip et al., 2013).

To get a more conclusive answer to the question whether RTs can pick up differences between lying and truth telling, a meta-analytic synthesis is warranted. A combination of data from multiple studies can overcome the restrictions and peculiarities of single studies and provides a more comprehensive view on the size and robustness of the RT deception effect (Chan & Arvey, 2012). Three meta-analyses have been published on this topic. Two large meta-analyses focused upon verbal, nonverbal, and para-verbal cues of deception during interview situations. The RT effects in these two meta-analyses were small and nonsignificant (DePaulo et al., 2003; Zuckerman et al., 1981; d = 0.02, 95% CI [−0.06; 0.10] and d = 0.13, respectively[2]). A possible reason for these disappointing results is that in the included studies, RTs were recorded under suboptimal conditions. This point was raised by Verschuere, Suchotzki, and Debey (2015), who argued that meaningful RT measurement must meet certain criteria. First, it needs to

[2] Cohen's d for independent samples was used in those studies.

be precise (i.e., computer-based). Second, participants should be able to respond immediately after stimulus presentation and be instructed to respond as fast as possible. Third, RTs need to be averaged over multiple measurements, with a proposed minimum of about 20 measurements per condition (i.e., lying vs. truth telling). Unfortunately, none of the studies that were included in the two meta-analyses mentioned above met these criteria: RTs were often assessed using stop-watches, participants were not prompted to respond as fast as possible but responded within the context of natural interview situations, and most studies employed relatively brief interactions with a very limited number of measurements. In more recent years, new computerized deception paradigms have been developed that overcome those limitations.

The third meta-analysis reviewed studies using a computerized RT measure (Agosta & Sartori, 2013). This meta-analysis focused exclusively on the autobiographical Implicit Association Test (aIAT; see method section), and revealed a significant medium-sized average effect (D-IAT = 0.58, 95% CI [0.41; 0.73], $k = 17$; $n = 412$; with the D-IAT measure roughly corresponding to Cohen's d, see Greenwald, Nosek, & Banaji, 2003). However, this meta-analysis was restricted to one particular paradigm. Furthermore, the analysis included studies in which participants were not aware of the deception detection context and were not instructed to deceive. This limits the generalizability and validity of the findings.

The Current Meta-Analysis

The current meta-analysis focuses on studies that report results of computerized RT measures from different RT deception paradigms. In doing so, we overcome what is known as the mono-measure bias (e.g., Smith, Fischer, & Fister, 2003). By using several deception measures instead of focusing on one, we aimed to comprehensively capture different aspects of lying while at the same time avoiding capturing effects that may only be specific to one paradigm. We combined results of RT deception paradigms that allow a within-subject comparison of truth telling and lying, resulting in estimates of the size, the precision, and the robustness of the RT deception effect. As we expected systematic differences between studies, we also investigated potential factors that may moderate the RT deception effects. In particular, we were interested in the moderating effects of variables that have been proposed to influence the cognitive load of lying, or may even reverse effects—indicating that in some situations truth telling may be cognitively more demanding than lying (Walczyk et al., 2014).

Publication Bias

Publication bias, that is, the tendency that large and significant effects are more likely to be published than small and nonsignificant effects, poses a problem in many research areas and may lead to overestimations of effect sizes in meta-analyses (e.g., Fanelli, 2012; Rosenthal, 1979; Sterling, 1959; Sterling, Rosenbaum, & Weinkam, 1995). We aimed to detect and, if necessary, correct our average effect size for the presence of publication bias using standard publication bias analyses (Borenstein, Hedges, Higgins, & Rothstein, 2011; Duval & Tweedie, 2000; Egger, Smith, Schneider, & Minder, 1997). In addition, two other indicators of publication bias were taken into account.[3] First, we investigated whether there are differences between studies that employ **RT as a primary versus secondary measure**. For example, fMRI studies with significant effects in BOLD responses may be published more easily, even if RT effects are not significant. In case of the presence of publication bias, we therefore expected larger effects in studies reporting RT as a primary measure compared with studies reporting RT as a secondary measure. Second, we investigated the effect of **publication year** on the effect sizes. In many research areas, one can observe a decline in effect sizes over the years. This so-called "decline effect" (Schooler, 2011) or "Law of initial results" (Ioannidis, 2005) was found in a recent meta-analysis by Meijer, Klein Selle, Elber, and Ben-Shakhar (2014) on autonomic nervous system (ANS) and event-related brain potential (ERP) measures of concealed information.[4] Because of a higher likelihood of publication bias, the decline effect may be particularly present in studies that used RT as a primary measure as compared with studies that used RT as a secondary measure.

Moderator Analyses

We wanted to explore the influence of potential moderator variables that can be related to the internal validity of the employed RT measures. Verschuere et al. (2015) proposed that to estimate the minimal time people need to initiate a truthful or deceptive response, participants must be instructed to respond as fast as possible. Hence, we coded whether **speed instructions** were provided to participants, and expected that studies in which speed instructions were given would yield larger effect sizes compared with studies in which speed instructions were not given.

We also wanted to assess the influence of the **stimulus presentation pace**. On the one hand, slower pace may enable participants to better control their performance and thereby result in smaller RT deception effects. On the other hand, two studies that have experimentally manipulated the response-stimulus interval in a RT-based deception paradigm revealed larger RT deception effects for a slow compared with a fast stimulus presentation, possibly because slower stimulus presentation makes it more difficult for participants to focus on the task goal, that is, deception (Debey et al., 2012). Based on the dual-process theory of Kane and Engle (2003) and Engle and Kane (2004), Debey et al. (2012) argued that their findings indicated that two mechanisms contributed to the RT deception effect: The time consuming nature of the resolution of the conflict between truth telling and lying, and occasional lapses in goal-maintenance.

A further moderator relates to the extent to which participants **attended to the stimuli** (Suchotzki, Verschuere, Crombez, & De Houwer, 2013). Some researchers designed their tasks in such a

[3] We also investigated whether there would be a difference between **published versus unpublished studies**. If a publication bias is present, one may expect unpublished studies to produce smaller effects than published studies. Results revealed no significant difference, $Q(1) = 1.961$, $p = .161$, yet this result should be treated with caution due to the very small sample of unpublished studies.

[4] It is important to note here that unlike the RT-Concealed Information Test (CIT; see method section), which mostly involves deception because participants have to deny knowledge of to-be-concealed items, deception is not always required in the ANS-based CIT. Although a comparison of these methods is still interesting as they belong to the same research field and serve the same purpose, this restricts the predictions we can draw from results with the ANS-based CIT for the RT CIT and deception in general.

way that attention to the stimulus content is guaranteed (e.g., by using targets that required a special response, memory tests, catch trials, or paradigms requiring alternating responses). We expected those studies to result in larger effects compared with studies that did not guarantee attention to the stimulus content.

In addition to moderators related to internal validity, we investigated a number of factors that are relevant from a theoretical or applied perspective. These factors have been proposed to influence or even reverse the cognitive cost of deception (DePaulo et al., 2003; Verschuere & De Houwer, 2011; Walczyk et al., 2014). It has been argued that **the type of paradigm** may affect the RT deception effect (Suchotzki et al., 2013; Verschuere & De Houwer, 2011). Several paradigms are available that allow a within-subject comparison between truthful and deceitful responses (i.e., the Concealed Information Test, the autobiographical Implicit Association Test, the Sheffield Lie Test, and the Differentiation of Deception Paradigm; see method section for detailed descriptions). We also investigated the influence of the total **number of trials**. From a psychometric point of view, one would expect more trials to be associated with enhanced reliability (Kleinberg & Verschuere, 2015). At the same time, more trials may lead to adaptation or habituation (Thompson, 2009) and may provide participants with more opportunities to practice their responses (DePaulo et al., 2003; Walczyk et al., 2014) or even to exert countermeasures (i.e., strategies that participants use to appear truthful; see, e.g., Ben-Shakhar, 2011). Previous findings regarding the effect of practice on the RT deception effect are inconclusive (Johnson, Barnhardt, & Zhu, 2005; Vendemia, Buzan, & Green, 2005). Research has also indicated that the **proportion of truth and lie trials** may be important. Larger proportions of truth trials are associated with larger RT deception effects in the Sheffield Lie Test (see method section; Van Bockstaele et al., 2012, 2015; Verschuere et al., 2011). Comparable results have been found with ANS and ERP concealed information measures (Ben-Shakhar, 1977; Ben-Shakhar, Lieblich, & Kugelmass, 1982; Hu, Hegeman, Landry, & Rosenfeld, 2012; Lieblich, Kugelmass, & Ben-Shakhar, 1970).

With regard to the **motivation to deceive**, there are two opposing hypotheses. The motivation impairment hypothesis (DePaulo, Kirkendol, Tang, & O'Brien, 1988) predicts that stronger motivation to avoid detection paradoxically leads to larger RT deception effects. This prediction was supported by the results of earlier deception and concealed information meta-analyses, which found larger effects when participants received motivational instructions or an extra incentive to successfully avoid detection (Ben-Shakhar & Elaad, 2003; DePaulo et al., 2003; Meijer et al., 2014). However, the motivation impairment effect has not been found in all concealed information measures (e.g., using ERPs, Ellwanger et al., 1996; using RTs, Kleinberg & Verschuere, 2015) and has been challenged (Burgoon & Floyd, 2000). In RT deception paradigms, stronger motivation to deceive might lead participants to assign more resources to cognitive control (Walczyk et al., 2014). Stronger motivation may further lead to more and better attempts of participants to employ countermeasures and fake a truthful test outcome by slowing down their responses on truth trials. Additionally, as the **relevance** of the information about which participants lie is closely associated with participants' motivation, the relevance of information can be expected to influence the RT effect size. Numerous studies have shown larger CIT effects for relevant compared with less relevant information (e.g., Abe et al., 2006; Carmel, Dayan, Naveh, Raveh, & Ben-Shakhar, 2003; Gamer, Kosiol, & Vossel, 2010; Jokinen, Santtila, Ravaja, & Puttonen, 2006; Lieblich, Ben-Shakhar, & Kugelmass, 1976) and the same effect could be expected in RT paradigms. Yet, one could also expect information relevance to be related to participants' efforts during the task.

To enhance external validity, it is important for deception detection methods to be tested in **populations** that are similar to populations in which those methods may finally be used (i.e., forensic populations). We therefore aimed to examine whether there are differences in the size of the RT deception effect depending on the type of population from which the study sample was selected. Specifically, we wanted to compare three types of sample populations: Normal/student populations, forensic populations, and clinical populations. Finally, to provide an indication about the use of RT measures in applied deception contexts, we aimed to explore the vulnerability of RT deception measures to strategic manipulations (i.e., countermeasures). In **countermeasure** studies, participants usually receive information about the test principle, the data pattern that would indicate deceit, measures to counteract this pattern, and may get hands-on practice to implement those measures to reduce the RT deception effect. As the inclusion of such studies would likely mask the cognitive dynamics of lying, we performed an additional meta-analysis on countermeasures studies. As summarized by Ben-Shakhar (2011), nearly all known concealed and deception detection methods are vulnerable to countermeasures, which is why we also expected a considerably reduced effect.

In sum, the current study provides a meta-analytic assessment of the RT signature of deception in computerized paradigms: It examines whether lies can be discerned from truth within the same participants, and whether factors related to publication bias, internal, external, and procedural validity moderate the cognitive cost of deception and hence influence the RT deception effect.

Method

Literature Search

Studies were identified through a search of the electronic databases PsychLIT, Web of Science, and PubMed using the following search string: ("Guilty knowledge" OR "Concealed information" OR GKT OR CIT OR "guilty action" OR "memory detection" OR "Concealed knowledge" OR "decept*" OR "lie detection" OR "lie-detection" OR "decei*" OR "Sheffield lie test" OR "autobiographical IAT" OR "aIAT" OR "autobiographical implicit association test") AND ("Reaction time*" OR "Response time*" OR "Response latenc*" OR ERP OR "Event related potentials" OR fMRI OR P300 OR scanner OR brain OR "neuro*"). In addition, the references of relevant review articles (Verschuere & De Houwer, 2011; Verschuere et al., 2015) and meta-analyses (Agosta & Sartori, 2013; DePaulo et al., 2003; Zuckerman et al., 1981) were searched for missing references, and the final list was checked by the coauthors.

Inclusion Criteria

The following criteria were used to select studies for inclusion in the meta-analysis:

1. The study was an experimental study reporting original data, and the full research report was published in English before July 2015 in a peer-reviewed journal. During an additional search, completed yet unpublished PhD theses were included. This search is described in more detail below.

2. The study sample consisted of adults or the mean age of the sample was at least 18 years.

3. The study reported results of a computer-based RT measure of deception. We also included studies that focused on other primary measures (e.g., ERPs, fMRI). Studies were excluded if no RT results were reported, or if RT was measured in suboptimal conditions. Examples for suboptimal conditions are designs that allowed participants to prepare their responses during a delay between presentation of the crucial stimulus and the response (e.g., Ambach, Stark, Peper, & Vaitl, 2008a), or if the experimenter rather than the participant pushed the response buttons (e.g., Abe et al., 2006).

4. Studies emphasized deception or information concealment through instructions.

5. The study enabled a within-subject comparison of deceptive and truthful responses. For instance, studies were excluded if it was not specified (either by the experimenter or by the design) on which trials participants were supposed to lie or tell the truth (e.g., Greene & Paxton, 2009; van Hooff, Sargeant, Foster, & Schmand, 2009).

6. The conditions included in the meta-analysis (deceptive vs. truthful responses) contained at least 20 presented trials, as estimates based on RT measures can be unreliable when the number of observations is limited (Verschuere et al., 2015).

7. The samples consisted of a minimum of 10 participants per condition. It is known that studies with small sample sizes are more susceptible to publication bias and tend to overestimate effect sizes (Sterne, Gavaghan, & Egger, 2000; Turner, Bird, & Higgins, 2013; Zhang, Xu, & Ni, 2013).

8. Conditions involving the use of rTMS, drugs, or alcohol administration, or instructions regarding the use of countermeasures were excluded. Such manipulations can be expected to mask the effect. In most of these cases, there were control conditions available that were included in our sample (e.g., Verschuere, Schuhmann, & Sack, 2012). Although not included in the main sample, studies instructing countermeasure use were included in the additional countermeasure meta-analysis.

9. Sufficient data for the computation of effect sizes had to be available. If a study reported RT results but did not contain sufficient information to calculate the effect size, the authors were contacted to provide the missing data. If this data could not be provided, the study was excluded.

The electronic databases were searched for references on the 18th and 19th of November 2013. To update our sample, an additional search was conducted on the 8th and the 9th of July 2015. Together, after the removal of duplicates, these searches resulted in 2003 references. A two-step procedure was used. In a first step, one reviewer (the third and the first authors respectively for the first and the second searches) screened the abstracts of all references. The reviewer was not blind to authors, institutions, journals, and results. If there was any doubt about exclusion, the references were retained for the full text search. In this first step, 1653 references were excluded. The main reasons for exclusion were that these references were not peer-reviewed research articles, were irrelevant/unrelated to deception, did not employ computer-based measures, or did not report original data. In the second step, full copies of the remaining 350 references were obtained and read. Where there was doubt (40 references), a second reviewer (the first author and the third author for the first and the second search, respectively) also read the article. If no consensus could be reached, a third reviewer was consulted (the second or the last author). In 20 cases, an e-mail was sent to the authors to ask for clarification on design features. Seventy publications were considered eligible for the analysis. The reasons for the exclusion of the 281 references from the full text search were as follows: (a) no original data (24 references), (b) participants were not required to deceive (76 references), (c) no within-subject comparison between lying and truth telling (42 references), (d) no computer-based RT paradigm and/or no report of RT results (106 references), (e) the study presented fewer than 20 trials per condition (19 references), (f) participants were taught countermeasures (1 reference), (g) results were based on a sample of less than 10 participants per group (10 references), (h) authors were asked for clarification, but did not reply to our request after two reminder e-mails (3 references). Note that as soon as the reviewer found one violation of an inclusion criterion, screening for other violations was stopped.

The searches in previous reviews and through the authors resulted in 7 additional publications that fulfilled our criteria. To further limit publication bias, a search for unpublished PhD theses was performed on the 11th of July 2015. We searched the electronic database Dissertation Abstracts, using the same search terms and inclusion criteria as outlined above (criteria 2 – 9). Additionally, we contacted the authors that contributed data to our sample and authors who may have supervised PhD research on RT deception studies. We also sent a call for PhD theses containing unpublished RT deception data via Twitter and announced our search at the "Conference for Psychology and Law" (Nuremberg, 2015) and the "Decepticon Conference" (Cambridge, 2015), both of which are international conferences that are visited by deception researchers. This search resulted in 33 PhD theses, many of which were already published. Two PhD theses fulfilled our inclusion criteria and contained unpublished data from 8 independent studies.

Taken together, the electronic and additional searches resulted in 79 eligible publications. To obtain all data necessary for the calculation of the effect sizes, the authors of 24 publications were contacted with the request to provide missing data. Most authors responded to this request. Six studies were excluded because the authors failed to provide clarifications on inconsistencies, or failed to send their data after two reminder e-mails (3 studies), or the authors stated that the data were no longer available (3 studies). Our final sample consisted of 73 publications, reporting a total of 115 independent studies. These numbers refer to our main meta-analysis; the studies included in our additional countermeasure meta-analysis and further details on that meta-analysis are described in the result section.

Coding System and Coding Decisions

We used a standard coding system to code every study in terms of sample and design characteristics. The coding categories were developed a priori, but if necessary adapted in an iterative process.

For sample characteristics, we coded year of publication, sample size (*n*), mean age of the sample, percentage of female participants, and whether participants were from a normal/student population, a forensic population, a clinical population, or a mixed forensic-clinical population. We coded whether RTs were the primary or the secondary outcome measure of the study and whether a study was published or unpublished. For design characteristics, we coded whether participants were instructed to respond as fast as possible (No vs. Yes),[5] the number of trials on which participants were instructed to lie, the number of trials on which participants were instructed to tell the truth, the absolute number of trials in the task, and the proportion of lie and truth trials (more lie trials, equal proportion, more truth trials). We also coded the intertrial interval—the interval from the beginning of one trial to the beginning of the next trial ($<$4000 ms, 4000–8000 ms, $>$ 8000 ms). For cases in which the intertrial interval was dependent on the response speed, we used the mean RT of truth telling and lying in that particular study as an estimate. Furthermore, we coded the category of the critical information (Real crime; Mock crime; Autobiographical; Other),[6] and its relevance [High (e.g., real crime, central mock crime details, or important autobiographical information such as first name, birth date etc.) versus Low (e.g., peripheral mock crime information, trivial autobiographical information such as favorite color, trivial semantic information, or other trivial information)]. We coded the type of deception paradigm (Concealed Information Test, autobiographical Implicit Association Test, Sheffield Lie Test, or Differentiation of Deception Paradigm) and whether the task could be performed without paying attention to stimulus content (No vs. Yes). We also coded whether it was reported that the participants were given extra motivational instructions to lie [None, Motivational instructions (e.g., "try to beat the test", "lie as convincingly as possible"), or an Incentive (e.g., extra money, extra credits) on top of the motivational instructions].

Some additional data-extraction and coding decisions were made. First, when several test repetitions were administered without manipulating any of our moderating factors, we used the data that were collapsed across the different repetitions. Second, if a study reported sufficient data to calculate the effect sizes for subgroups or within-subject conditions (with manipulations that were not of interest for the current meta-analysis), but did not report the collapsed results, we coded the subgroups or different conditions separately (as independent or dependent samples, respectively). However, if the reported data were insufficient to calculate either effect size, we requested the authors to send us the data averaged over the respective groups or conditions. Authors were contacted to provide the means and standard deviations for lying and truth telling and the correlation (Pearson's *r*) between these two conditions.

Data-extraction and coding was conducted by one reviewer (the first author) using an excel sheet specifically designed for this meta-analysis. When in doubt, a second or third reviewer (the second or last author) was asked to resolve disagreements. The final coding reflects the consensus of the coding.

Paradigms

We only included paradigms that allowed a within-subject comparison of truth telling and lying. All studies that were included were categorized as one of four different paradigms: The Concealed Information Test, the autobiographical Implicit Association Test, the Sheffield Lie Test, or the Differentiation of Deception Paradigm.

The **Concealed Information Test** (**CIT**; Lykken, 1959) is a method designed to measure recognition of critical (e.g., crime-relevant) information. In CIT experiments, participants are usually instructed to hide particular knowledge (e.g., about crime-related facts, autobiographical details, or a card picked from a card set) from the experimenter. During the CIT, participants are presented with stimuli referring to this knowledge, embedded among well-matched irrelevant stimuli. Crucially, in most RT-based CITs (Seymour et al., 2000), participants are instructed to deny knowledge of both types of stimuli (i.e., respond "no"), meaning that they lie on trials on which critical stimuli are presented and tell the truth on trials on which irrelevant stimuli are presented. In most of these CIT versions, a third type of stimuli is also used. These so-called target stimuli are mostly learned just before the administration of the CIT and participants are instructed to admit their recognition (i.e., respond "yes"). These stimuli guarantee that participants have to process the content of the stimuli and cannot perform the task by simply responding "no" to every single stimulus. Although target stimuli also require truthful responses, they are not taken into account for the calculation of the CIT-effect ($RT_{critical\ stimuli} - RT_{irrelevant\ stimuli}$) and for the lie–truth contrast in our meta-analysis.

[5] Initially, we planned to code studies in which no such instructions were reported as "No." Because too many studies did not report any instructions, we decided to contact the authors in all those cases. With a very high response rate, we obtained the information for all but three of the studies (see also result section).

[6] We originally planned to analyze the influence of stimulus category, as this was a significant moderator in two meta-analyses on the autonomic CIT effect (Ben-Shakhar & Elaad, 2003; Meijer et al., 2014). With stimuli coded as real crime ($k = 0$), mock crime ($k = 17$), autobiographical ($k = 50$), or other ($k = 47$), there was no significant difference between categories, $Q(2) = 0.219$, $p = .896$. We noticed, however, that an interpretation of this finding is difficult because of the large number of studies in the "other" category. Unfortunately, because of the large variety of paradigms and stimuli used, it proved difficult to design a narrower, yet mutually exclusive subdivision of this "other" category. We therefore decided to exclude this analysis from the current report. Excluding the "other" category and only comparing the mock crime and autobiographical category also failed to reveal a significant difference $Q(1) = 0.215$, $p = .643$. We also coded whether responses had to be given with one or two hands (One vs. Two). This information was only available for 33 studies and results revealed no significant differences between both types of responses, $Q(1) = 0.444$, $p = .505$.

The autobiographical Implicit Association Test (aIAT; Sartori, Agosta, Zogmaister, Ferrara, & Castiello, 2008) is a method designed to reveal which of two contrasting autobiographical events is true. In the aIAT, participants are asked to classify four different types of stimuli using two response keys. The four different types of stimuli are true statements (e.g., "I'm currently doing a computerized test"), false statements (e.g., "I'm currently reading a crime novel"), and statements related to two contrasting autobiographical events (e.g., "I robbed the bank" vs. "I was at the movies"). These statements have to be classified into four categories, TRUE versus FALSE, and for instance in the previously used example, I ROBBED THE BANK versus I WAS AT THE MOVIES. Crucially, the aIAT always consists of two test blocks during which different categories share the same response keys. In one block, true statements and statements relating to one autobiographical event are mapped to one response key and false statements and statements relating to the other autobiographical event are mapped to the other key. In the other block, the mapping is reversed. The idea behind the aIAT is that the FALSE and TRUE categories interfere with the categorization of the statements referring to the two autobiographical events under investigation. Thereby, participants are indirectly deceiving when categorizing the autobiographical true event (e.g., the robbery) into the FALSE category and the autobiographical false event (e.g., going to the movie) into the TRUE category. In our meta-analysis, we included aIAT experiments in which participants were told about the deception context of the experiment and were explicitly instructed to hide or conceal their true autobiographical memory. We coded the block in which both types of true statements were mapped to the same response key and both types of false statements were mapped to the other response key as truthful control responses, and the block with the reversed mapping as deceptive responses.

The Sheffield Lie Test (SLT; Spence et al., 2001) is a paradigm in which participants are instructed to lie or tell the truth about a set of stimuli, depending on a cue (e.g., lie to questions in blue, tell the truth to questions in yellow). Importantly, in the SLT, participants tell the truth and lie about the same set of stimuli, which served as the crucial contrast included in our meta-analysis.

The remaining paradigms that contrasted lying and truth telling on different items were variants of the Differentiation of Deception Paradigm (DoD; Furedy et al., 1988). In the DoD, unlike in the SLT, stimuli do not form their own control. Thus, participants are cued or instructed to lie about one stimulus set, and to tell the truth about another stimulus set. One example is the old/new word paradigm, in which participants categorize words as old (i.e., already seen) and new (i.e., not yet encountered). In the deceptive condition, participants are instructed to lie (i.e., say "old" for a subset of "new" words and say "new" for a subset of "old" words), whereas in the truth condition they are instructed to respond truthfully. For the DoD, the crucial contrast that we included is the comparison between the deceptive and the most adequately matched truthful responses (technically one also tells the truth to filler or catch trials that are used to assure attention, but such trials were excluded as they differ markedly from the deceptive responses).

Meta-Analytic Procedures

All our analyses and computations were carried out using Comprehensive Meta-Analysis software, Version 2.2.050 (Biostat, Englewood, NJ) and our meta-analytic approach is based on Borenstein et al. (2011).The effect size used in our meta-analysis was the standardized paired difference (Cohen's d for paired data; Borenstein et al., 2011; Dunlap, Cortina, Vaslow, & Burke, 1996; Morris & DeShon, 2002). As a rule of thumb, Cohen (1988) proposed 0.20, 0.50, and 0.80 as thresholds for "small", "moderate", and "large" effects, respectively. The standardized paired difference was calculated by subtracting the mean RT for truth telling from the mean RT for lying. This difference was then divided by the pooled standard deviation. The pooled standard deviation was corrected for the correlation between lying and truth telling ($SD_{\text{pooled}} = \sqrt{SD^2_{\text{lie}} + SD^2_{\text{truth}} - 2 * r_{lie,truth} * SD_{\text{lie}} * SD_{\text{truth}}}$). If the standard error was given instead of the standard deviation, the following formula was used $SD = SE * \sqrt{n}$.

When the correlation was not available or could not be derived from the results, we calculated and imputed the mean correlation between lying and truth telling across all studies for which correlations were available ($k = 36$; $n = 766$) using a meta-analytic approach. Cochran's Q indicated a significant heterogeneity between correlations, $Q(35) = 134.597$, $p < .001$ ($I^2 = 74\%$). Further analyses revealed that this heterogeneity in correlations was partly explained by differences between paradigms, $Q(3) = 10.565$, $p = .014$. Post hoc comparisons revealed significant differences between the correlations of the aIAT studies and all studies using other paradigms (all ps $< .02$). There were no significant differences between the mean correlations in the other three paradigms (all ps $> .80$). We therefore decided to calculate and impute the correlations for aIAT studies separately, $r = .698$ (95% Confidence Interval (95% CI) [.527; .815], $k = 3$; $n = 78$). There was no aIAT study in which a correlation needed to be imputed. The imputed correlation for all other paradigms was $r = .880$ (95% CI [.838; .911], $k = 33$; $n = 688$).

When the means and the standard deviations for both conditions were not available, but the t values for the crucial contrast (lying vs. truth telling) were reported, d was derived from this t value ($k = 2$; $n = 73$; Morris & DeShon, 2002). If only an F value for the main effect of lying versus truth telling was given, the t value was calculated as $t = \sqrt{F}$ and d was derived from this t-value ($k = 22$; $n = 551$). Standardized paired differences with a positive sign (+) indicate effects of longer RTs for lying compared with truth telling, whereas effect sizes with a negative sign (−) indicate effects in the opposite direction.

In our meta-analysis, we took into account the sampling error of each sample. Effect sizes were weighted by the inverse of the estimated sampling variance, whereby effect sizes derived from larger samples gain more weight than effect sizes derived from smaller samples (Borenstein et al., 2011). For the calculation of the average effect size across all moderators, we chose a random effects model. When reporting effect sizes (d), we always provide the 95% CIs, the number of independent studies (k), and the total number of participants (n). We applied Cochran's Q test to judge the degree of heterogeneity in effect sizes (Borenstein et al., 2011). We also provide I^2, which denotes the percentage of observed variance that is due to real differences between studies and as such may be explained by moderating variables. According to Higgins, Thompson, Deeks, and Altman (2003), 25%, 50%, and 75% could be considered as low, moderate, and high scores for I^2, respectively.

Publication bias was assessed using a funnel plot, which depicts the effect size of each study on the *x* axis versus its precision (e.g., the inverse of the standard error) on the *y* axis. The rationale for a funnel plot is that if a meta-analysis captured all relevant studies, the dispersion of the studies is expected to be symmetrical. Asymmetry, especially at the bottom of the plot, is taken as an indication of missing studies and therefore as an indication of a publication bias (e.g., Borenstein et al., 2011) or a file drawer problem (Rosenthal, 1979). The presence of a publication bias usually leads to an overestimation of the mean effect size. To get an estimate of publication bias, we first statistically evaluated the relationship between effect size and sample size with the Egger test (Egger et al., 1997), a regression analysis predicting the standardized effects (effect size/standard error) from precision (1/standard error). When there was a significant result (i.e., a significant intercept), we used the Trim and Fill method (Duval & Tweedie, 2000) to impute the estimated effect sizes of missing studies and then recompute the average effect size. This provides an estimate of the average effect size had the funnel plot been symmetrical.

To address whether variations in effect sizes can be explained by categorically coded variables, we performed moderator analyses using a mixed effects model. Such analyses can be considered as a meta-analytic ANOVA analogue, as the variance in effect sizes is partitioned into the portion explained by the categorical variable (Q_B) as an indicator of variability between group means, and the residual portion (Q_W) as an indicator of variability within groups. Q_B is tested for significance against a chi-square distribution with $df = k - 1$. A significant between-groups effect indicates that the variance in effect sizes is at least partially explained by the moderator variable. Moderator levels with five or less studies were not included in these analyses. If the moderator analysis involved more than two valid groups, simple contrasts were applied to determine significant differences between groups. Note that this significance test is not identical to the test of overlap of confidence intervals around the respective means. That is, even in the case of confidence interval overlap, between-groups differences can be statistically significant (Cumming & Finch, 2005; Estes, 1997). This is because the between-groups test is based on a joint estimate of the standard error of the difference between means, whereas confidence intervals around a single mean do not reflect between-groups information. In that context, the confidence intervals should be best interpreted as providing information about the precision of single effect size estimates and their difference from zero (Rouder & Morey, 2005).

To maintain the independence of the effect sizes used in our meta-analysis (Lipsey & Wilson, 2001), whenever necessary, we averaged effect sizes across within-subject conditions (Hedges & Olkin, 1985). When this turned out to be impossible, such as for a moderator analysis including studies with two within-subject conditions that differed on the respective moderator variable, we only used the data from the moderator category that contained the least studies for that respective analysis to increase power in that particular category (see Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & Van Ijzendoorn, 2007; Crombez, Van Ryckeghem, Eccleston, & Van Damme, 2013). Finally, for the continuously coded variables (i.e., the publication year and the absolute number of trials), we performed metaregressions using the methods of moments procedure (Thompson & Higgins, 2002). The outcome of this metaregression is reported as a point estimate of the slope and the significance of that slope.

Seventy-three publications reporting 115 independent studies were initially included in the meta-analysis. Screening the data set for outliers revealed 1 study from 1 publication with an effect size that strongly deviated from the average effect size (>125 *SD*s; $d = -6.912$; 95% CI [−8.829; −4.994] for Langleben et al., 2005). Closer inspection revealed that in this study, a paradigm was used in which the crucial comparison of lying versus truth telling was confounded with the ratio of yes versus no responses. Yes responses were infrequent (1:10), and only used for truth trials, whereas No responses were frequent (9:10) and used for lie and filler trials. As the effect size of this study was an outlier, we decided to exclude it from further analyses. As a consequence, our meta-analysis is based on a total of 114 independent studies reported in 72 publications.

Results

Study Characteristics

Of the 114 included studies, 34 used the CIT, 9 used the aIAT, 55 used the SLT, and 16 used variants of the DoD. All but one study were published after 2000, indicating the increased use and report of computer-based RT as a primary or secondary measure of deception. The distribution of all coded categorical variables can be found in Table 2.

Based on the studies for which the RT means were available ($k = 90$), the overall average RT difference between lying and truth telling was 115 ms, 95% CI [103; 126]. More specifically, it was 49 ms, 95% CI [41; 57] for the CIT ($k = 32$), 149 ms, 95% CI [53; 245] for the aIAT ($k = 3$), 180 ms, 95% CI [151; 209] for the SLT ($k = 47$), and 106 ms, 95% CI [67; 145] for the DoD ($k = 8$).

Overall Effect Size

We computed the average standardized paired difference across all studies ($k = 114$; $n = 3307$). The results revealed a large and significant overall average effect size, $d = 1.256$, 95% CI [1.137; 1.374]. The effect sizes of the single studies and the average effect size across all studies are provided in Table 1. Caution is warranted when interpreting the average effect size, as Cochran's Q revealed large heterogeneity in the sample of studies, $Q(113) = 697.638$, $p < .001$, justifying our a priori chosen random effects model. I^2 indicated that about 84% of the observed variance between effect sizes was caused by systematic differences between studies, and this warranted the search for potential moderators of the RT deception effect.

Publication Bias Analysis

Publication bias was assessed with the funnel plot depicted in Figure 1. Graphical inspection of the plot revealed a slight asymmetry, especially at the bottom of the plot. This observation was supported by the significant result of the Egger test, with the value of the intercept = 2.747, $p < .001$. The trim and fill method revealed the need to impute 23 missing studies to eliminate this asymmetry. Recalculating the average effect size after the imputation of the missing studies revealed a lower, but still large and

Table 1
Summary of the Studies Included in the Meta-Analysis and the Average Effect Size Across Studies

Study	Paradigm	*n*	*d*	95% CI *LL*	95% CI *UL*
Abe et al. (2008)	DoD	20	.870	.355	1.385
Agosta et al. (2011) - Experiment 1	aIAT	14	1.400	.663	2.137
Agosta et al. (2011) - Experiment 2	aIAT	10	.394	−.250	1.037
Agosta et al. (2011) - Experiment 3	aIAT	18	1.282	.659	1.906
Agosta et al. (2011) - Experiment 4	aIAT	12	.207	−.365	.779
Allen, Iacono, & Danielson (1992)	CIT	20	2.012	1.250	2.774
Carrión, Keenan, & Sebanz (2010)	SLT	11	.675	.020	1.329
Cutmore et al. (2009)	CIT	17	.771	.199	1.343
Debey. Verschuere. & Crombez (2012) - Experiment 1 Depletion	SLT	33	1.016	.596	1.436
Debey. Verschuere. & Crombez (2012) - Experiment 1 Control	SLT	34	1.036	.619	1.452
Debey, Verschuere, & Crombez (2012) - Experiment 2 Depletion	SLT	23	1.375	.804	1.946
Debey, Verschuere, & Crombez (2012) - Experiment 2 Control	SLT	22	1.014	.499	1.528
Debey, De Houwer, & Verschuere (2014) - Experiment 1	SLT	26	1.558	.985	2.131
Debey, De Houwer, & Verschuere (2014) - Experiment 2–25% TD Group	SLT	24	1.873	1.209	2.537
Debey, De Houwer, & Verschuere (2014) - Experiment 2–75% TD Group	SLT	25	2.537	1.732	3.342
Debey et al. (2015) - Experiment 1	SLT	34	1.263	.812	1.713
Debey et al. (2015) - Experiment 2	SLT	51	1.967	1.497	2.437
Debey PhD Chapter 1 - Elderly	SLT	51	1.766	1.388	2.143
Debey PhD Chapter 1 - Middle adulthood	SLT	205	1.748	1.370	2.126
Debey PhD Chapter 1 - Older adulthood	SLT	129	1.823	1.366	2.280
Debey PhD Chapter 1 - Young adulthood	SLT	79	.758	.484	1.031
Debey PhD Chapter 3 - Experiment 1	SLT	54	2.197	1.704	2.689
Duran, Dale, & McNamara (2010)	DoD	25	1.796	1.163	2.430
Farrow et al. (2010)	SLT	40	1.085	.695	1.476
Farrow, et al. (2003)	SLT	61	.574	.303	.845
Fecteau et al. (2013)	SLT	11	1.607	.713	2.502
Fullam, McKie, & Dolan (2009)	SLT	24	.499	.074	.923
Gamer & Berti (2010) - Experiment 1	CIT	12	2.597	1.414	3.781
Gamer & Berti (2010) - Experiment 2	CIT	12	.366	−.219	.950
Gamer & Berti (2012)	CIT	20	1.478	.819	2.137
Gamer et al. (2007)	CIT	14	1.342	.620	2.064
Ganis & Schendan (2013)	CIT	17	1.609	.882	2.336
Ganis, Morris, & Kosslyn (2009)	DoD	14	1.020	.373	1.666
Ganis et al. (2011)	CIT	12	.917	.242	1.591
Hu, Wu, & Fu (2011)	SLT	22	2.697	1.797	3.596
Hu, Chen, & Fu (2012) - Control group	SLT	16	1.912	1.088	2.736
Hu, Chen, & Fu (2012) - Instruction group	SLT	16	2.164	1.268	3.060
Hu, Chen, & Fu (2012) - Training group	SLT	16	1.194	.553	1.836
Hu et al. (2013)	CIT	31	1.088	.644	1.532
Ito et al. (2011)	DoD	32	1.584	1.064	2.105
Ito et al. (2012)	DoD	16	1.605	.864	2.346
Jiang et al. (2013)	SLT	32	.967	.547	1.387
Johnson, Barnhardt, & Zhu (2003)	DoD	21	1.116	.571	1.661
Johnson et al. (2008)	SLT	16	1.011	.409	1.614
Kaylor-Hughes et al. (2011)	SLT	52	.310	.032	.589
Kleinberg & Verschuere (2015) - Experiment 1	CIT	88	1.174	.892	1.455
Kleinberg & Verschuere (2015) - Experiment 2	CIT	100	1.145	.882	1.408
Kubo & Nittono (2009)	CIT	18	.319	−.154	.793
Leue, Lange, & Beauducel (2012)	CIT	102	.335	.136	.534
Lui & Rosenfeld (2009) - Experiment 1	DoD	14	−.048	−.572	.476
Lui & Rosenfeld (2009) - Replication study	DoD	14	.815	.210	1.419
Lui & Rosenfeld (2009) - Experiment 2	DoD	13	−.089	−.634	.455
Mameli et al. (2013) - Patients with ET	SLT	15	1.029	.404	1.655
Mameli et al. (2013) - Patients with PD	SLT	20	1.595	.934	2.255
Mameli et al. (2013) - Healthy controls	SLT	17	1.569	.859	2.279
Marchewka et al. (2012)	DoD	29	1.891	1.282	2.500
Matsuda et al. (2009)	CIT	21	−.017	−.445	.411
Matsuda, Nittono, & Allen (2013)	CIT	19	.238	−.218	.694
Meijer et al. (2007)	CIT	24	1.905	1.233	2.576
Mertens & Allen (2008)	CIT	15	.926	.321	1.531
Noordraven & Verschuere (2013) Guilty group	CIT	21	1.646	.987	2.305
Noordraven & Verschuere (2013) Innocent group (aCIT)	CIT	21	2.033	1.284	2.782
Nose, Murai, & Taira (2009)	CIT	19	.318	−.143	.779

(*table continues*)

Table 1 (*continued*)

Study	Paradigm	*n*	*d*	95% CI *LL*	95% CI *UL*
Pfister, Foerster, & Kunde (2014) - Experiment 1	SLT	16	.430	−.082	.942
Pfister, Foerster, & Kunde (2014) - Experiment 2	SLT	16	.416	−.095	.926
Priori et al. (2008)	SLT	15	1.472	.742	2.203
Rosenfeld, Hu, & Pederson (2012)	CIT	10	1.275	.440	2.109
Sai et al. (2014)	CIT	16	3.616	2.271	4.962
Sartori et al. (2008) - Experiment 1	aIAT	37	1.004	.609	1.399
Sartori et al. (2008) - Experiment 4	aIAT	20	1.416	.796	2.036
Sheridan & Flowers (2010) - Experiment 1	DoD	13	2.157	1.166	3.149
Sheridan & Flowers (2010) - Experiment 2	DoD	12	1.779	.870	2.689
Sheridan & Flowers (2010) - Experiment 3	DoD	12	3.419	1.939	4.900
Spence et al. (2001) - Experiment 1	SLT	30	.630	.238	1.022
Spence et al. (2001) - Experiment 2	SLT	10	.727	.029	1.426
Suchotzki et al. (2013) - Experiment 1	SLT	25	1.507	.934	2.080
Suchotzki et al. (2013) - Experiment 2	SLT	21	.644	.174	1.114
Suchotzki, Crombez, Debey, et al. (2015)	SLT	88	.975	.721	1.229
Suchotzki, Crombez, Smulders, et al. (2015)	SLT	20	1.917	1.179	2.656
Suchotzki, Verschuere, et al. (2015)	CIT	32	1.034	.605	1.464
Van Bockstaele et al. (2012) - Frequent lie group	SLT	14	.520	−.038	1.078
Van Bockstaele et al. (2012) - Control group	SLT	14	1.154	.478	1.830
Van Bockstaele et al. (2012) - Frequent truth group	SLT	14	1.414	.673	2.155
Varga et al. (2015)	CIT	46	2.090	1.569	2.612
Vartanian, Kwantes, & Mandel (2012)	DoD	15	3.006	1.817	4.195
Vartanian et al. (2013)	DoD	15	1.915	1.063	2.767
Vendemia, Buzan, & Green (2005)	SLT	25	1.682	1.073	2.291
Vendemia, Buzan, & Simon-Dack (2005) - Experiment 3	SLT	38	1.408	.959	1.856
Verschuere et al. (2009)	CIT	16	1.129	.502	1.756
Verschuere, Prati, & De Houwer (2009) - Experiment 1	aIAT	18	.469	−.017	.956
Verschuere, Prati, & De Houwer (2009) - Experiment 2	aIAT	18	1.026	.455	1.597
Verschuere, Prati, & De Houwer (2009) - Experiment 3	aIAT	42	.494	.173	.814
Verschuere et al. (2010)	CIT	32	1.968	1.374	2.562
Verschuere et al. (2011) - Frequent truth group	SLT	21	2.580	1.690	3.470
Verschuere et al. (2011) - Frequent lie group	SLT	22	.356	−.075	.787
Verschuere et al. (2011) - Control group	SLT	21	.778	.290	1.266
Verschuere, Schuhmann, & Sack (2012)	SLT	26	1.278	.760	1.797
Verschuere, Kleinberg, & Theocharidou (2015) - Experiment 1 Multiple Probe	CIT	19	2.190	1.292	3.089
Verschuere, Kleinberg, & Theocharidou (2015) - Experiment 1 Single Probe	CIT	21	1.630	.839	2.420
Verschuere, Kleinberg, & Theocharidou (2015) - Experiment 2 Multiple Probe	CIT	44	1.701	1.239	2.163
Verschuere, Kleinberg, & Theocharidou (2015) - Experiment 2 Single Probe	CIT	53	.575	.284	.865
Visu-Petra, Miclea, & Visu-Petra (2012)	CIT	21	1.263	.690	1.837
Visu-Petra et al. (2013)	CIT	73	1.988	1.593	2.384
Visu-Petra et al. (2014)	CIT	47	2.207	1.677	2.737
Williams - PhD Experiment 6	SLT	18	1.682	.948	2.417
Williams - PhD Experiment 7	SLT	19	.857	.306	1.408
Williams - PhD Experiment 8	SLT	21	1.345	.751	1.940
Williams et al. (2013) - Experiment 1	SLT	19	.644	.150	1.139
Williams et al. (2013) - Experiment 2	SLT	22	1.962	1.247	2.676
Williams et al. (2013) - Experiment 3	SLT	36	1.971	1.410	2.531
Williams et al. (2013) - Experiment 4	SLT	32	1.913	1.330	2.495
Williams et al. (2013) - Experiment 5	SLT	30	2.202	1.540	2.865
Wu, Hu, & Fu (2009)	DoD	12	.579	−.034	1.192
Zhao, Zheng, & Zhao (2012)	CIT	16	1.785	.996	2.575
Average effect size (random effects model)		**3307**	**1.256**	**1.137**	**1.374**
Corrected average effect size			**1.049**	**.930**	**1.169**

Note. *n* = number of participants; *d* = standardized paired difference; 95 % CI = 95% Confidence Interval; *LL* = Lower Limit; *UL* = Upper Limit. Effect sizes used in our meta-analysis may differ from effect sizes reported in the respective articles, as we used the dependent Cohen's *d* and, if necessary, imputed the mean correlation (Pearson's *r*) derived from studies for which correlations were available.

significant estimate of the overall average effect size, $d = 1.049$; 95% CI [0.930; 1.169].

Moderator Analyses

Table 2 provides Cochran's Q for each of the categorical moderators, as well as the proportion of the variance of the RT deception effect that they account for. Furthermore, it provides the effect size and the Q statistic for each level of the variable, indicating whether there is still unexplained systematic variance left. Results of metaregressions (for continuous moderator variables) are reported in the text. A complete table showing the coding of all moderator variables for each study is available at https://osf.io/sphg3/#.

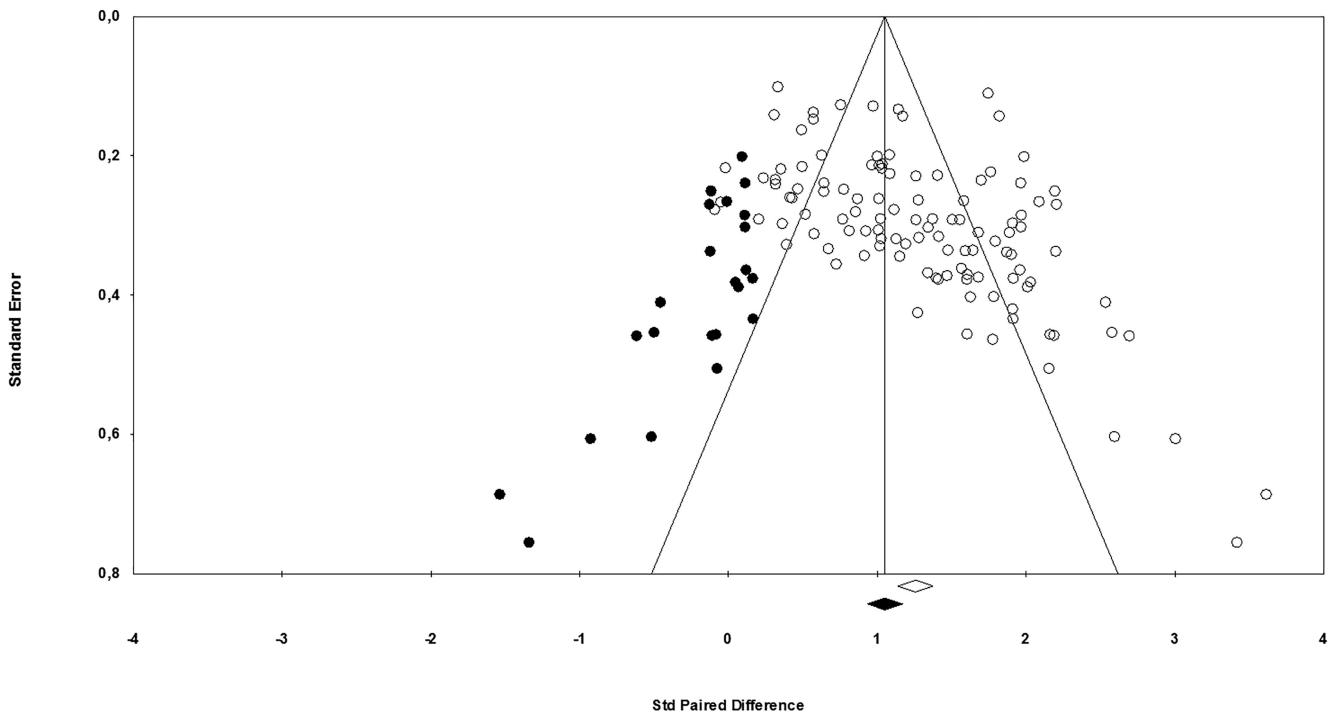

Figure 1. Funnel plot of effect size (Cohen's *d*) against standard error (plotted from high to low standard error). The unfilled circles indicate studies included in our meta-analysis and the filled circles indicate missing studies that were imputed using the trim and fill method. The unfilled diamond symbol indicates the estimated mean effect size and its 95% confidence interval without the imputed studies and the filled diamond symbols the estimated mean effect size and its 95% confidence interval with the imputed studies.

As an additional measure of publication bias, we investigated whether studies that reported RT as a primary measure resulted in higher effect sizes than studies that reported RT as a secondary measure, and whether there was an influence of publication year. Results revealed a larger mean effect size in studies that reported RT as a primary measure, $d = 1.354$, 95% CI [1.223; 1.486], compared with studies that measured RT as a secondary measure, $d = 1.064$, 95% CI [0.854; 1.274], $Q(1) = 5.271$, $p = .022$. This estimate is similar to the corrected *d* value obtained with the trim and fill method, thereby providing further evidence for a publication bias, which slightly inflated the average effect size estimate.

Excluding unpublished studies, a metaregression using the method of moments procedure revealed a significant effect of publication year, with the point estimate of the slope = 0.038, 95% CI [0.003; 0.073], $p = .032$. The significant slope of the metaregression line on the effect sizes indicates that effect sizes increased with $d = 0.038$ per (later) year of publication. Separating studies that reported RT as a primary measure and studies that reported RT as a secondary measure revealed that the significant effect of publication year was only present in the studies using RT as a primary measure (point estimate of slope = 0.057, 95% CI [0.016; 0.097], $p = .006$). The significant slope of the metaregression line for the effect sizes of the studies that reported RT as a primary measure indicates an increase in effect size of $d = 0.057$ per (later) year of publication within this subgroup. There was no significant effect of publication year in studies using RTs as secondary measure (point estimate of slope = -0.0004, 95% CI [-0.058; 0.058], $p = .989$).

Next, we investigated the influence of speed instructions, measures to ensure attention toward stimulus content, and population. For the calculation of the effect of speed instructions, 3 studies (from 2 publications) could not be included as the authors did not respond to our request for clarification ($k = 2$) or the information was no longer available ($k = 1$). The results indicated no significant difference between studies in which instructions to respond as fast as possible were given, $d = 1.285$, 95% CI [1.160; 1.410], compared with studies in which no such instructions were given, $d = 0.971$, 95% CI [0.642; 1.300], $Q(1) = 3.048$, $p = .081$. No moderator analysis could be run for whether a paradigm guaranteed that participants could not divert their attention from the stimuli, as there was only one study in which no such measures were taken (Matsuda et al., 2009). Similarly, it was impossible to assess the influence of the population that the participants belonged to, as there were only 2 studies with a clinical sample (Mameli et al., 2013), and only 2 studies with a combined clinical/forensic sample (Jiang et al., 2013; Kaylor-Hughes et al., 2011).

Next, we investigated a number of procedural moderators. Importantly, we found a significant overall effect of paradigm on the effect sizes, $Q(3) = 9.059$, $p = .029$. Post hoc comparisons revealed that the mean effect size for aIAT studies was with $d = 0.822$, 95% CI [0.538; 1.106] significantly smaller than the three other paradigms (all *p*s < .04). There were no significant differences between the other three paradigms (all *p*s > .78).

For the analysis of the effect of the inter trial interval, 21 studies (from 8 publications) were excluded because the information was not reported. Our results did not reveal a significant effect of the

Table 2
Results of the Categorical Moderator Analyses

Moderator	k	n	d + 95% CI	Q	I^2
RT prim/sec measure*	114	3307		$Q(1) = 5.271$, $p = .022$	
Primary measure	73	2442	1.354 [1.223; 1.486]	$Q(72) = 369.345$, $p < .001$	81
Secondary measure	41	865	1.064 [0.854; 1.274]	$Q(40) = 238.808$, $p < .001$	83
Speed instructions	111	3220		$Q(1) = 3.048$, $p = .081$	
No	11	354	0.971 [0.642; 1.300]	$Q(10) = 58.535$, $p < .001$	83
Yes	100	2866	1.285 [1.160; 1.410]	$Q(99) = 561.935$, $p < .001$	82
Paradigm*	114	3307		$Q(3) = 9.059$, $p = .029$	
CIT	34	1063	1.297 [1.060; 1.535]	$Q(33) = 253.484$, $p < .001$	87
aIAT	9	189	0.822 [0.538; 1.106]	$Q(8) = 21.043$, $p = .007$	62
SLT	55	1778	1.287 [1.129; 1.446]	$Q(54) = 312.853$, $p < .001$	83
DoD	16	277	1.350 [0.945; 1.755]	$Q(15) = 86.740$, $p < .001$	83
Inter trial interval	93	2893		$Q(2) = 2.563$, $p = .278$	
<4000 ms	56	2038	1.310 [1.133; 1.488]	$Q(55) = 443.593$, $p < .001$	88
4000–8000 ms	21	563	1.106 [0.884; 1.329]	$Q(20) = 86.169$, $p < .001$	77
>8000 ms	16	292	1.384 [1.017; 1.751]	$Q(15) = 72.223$, $p < .001$	79
Proportion lie/truth trials*[a]	106	3175		$Q(1) = 4.367$, $p = .037$	
More lie trials	2	36	—	—	—
Equal proportion	47	1653	1.441 [1.271; 1.612]	$Q(46) = 260.054$, $p < .001$	82
More truth trials	59	1522	1.184 [1.014; 1.355]	$Q(58) = 358.044$, $p < .001$	84
Stimulus relevance	114	3307		$Q(1) = 2.560$, $p = .110$	
Low	71	2617	1.210 [1.060; 1.361]	$Q(70) = 458.994$, $p < .001$	85
High	43	690	1.424 [1.209; 1.639]	$Q(42) = 271.889$, $p < .001$	85
Motivation*	114	3307		$Q(2) = 6.439$, $p = .040$	
None	85	2552	1.331 [1.189; 1.474]	$Q(84) = 541.480$, $p < .001$	84
Motivating instructions	23	626	1.002 [0.781; 1.223]	$Q(22) = 103.804$, $p < .001$	79
Incentive	6	129	1.132 [0.836; 1.427]	$Q(5) = 8.131$, $p = .149$	39

Note. k = number of independent studies; n = number of participants; d = standardized paired difference; 95% CI = 95% Confidence Interval, with 95% CIs that do not include zero also indicating that the respective effect is significant with $p < .05$; Q = Cochran's Q indicating for each of the categorical moderator variables whether the variable explains a significant portion of the variance of the deception effect sizes, and for each level of the variable whether there is unexplained systematic variance left; I^2 = percentage of observed remaining variance that is due to real differences between studies and as such may be explained by additional moderating variables; ms = milliseconds.
[a] Results reported for this factor are after the exclusion of the 2 studies using more lie trials.
* Significant with $p < .05$.

length of the intertrial interval, $Q(2) = 2.563$, $p = .278$. There was also no significant effect of the absolute number of trials on the effect sizes (point estimate of slope = 0.0003, 95% CI [−0.0004; 0.0010], $p = .366$). For the calculation of the influence of the proportion of lie trials on the effect sizes, we excluded 6 studies (from 4 publications) in which the proportion of lie trials could not be calculated. Furthermore, we could not include the studies belonging to the category "more lie trials" in our analysis, as there were only 2 studies available. Contrary to our prediction, our results revealed a larger average effect size in studies using an equal proportion, d = 1.441, 95% CI [1.271; 1.612], compared with studies using more truth trials, d = 1.184, 95% CI [1.014; 1.355], $Q(1) = 4.367$, $p = .037$.

Stimulus relevance did not produce a statistically significant effect $Q(1) = 2.560$, $p = .110$. It is possible that stimulus relevance is especially important in CIT designs, where relevant items are more likely to draw attention. We therefore ran an additional exploratory moderator analysis for stimulus relevance only for the studies using the CIT. Within the CIT studies, there was indeed a significant effect, $Q(1) = 11.327$, $p = .001$, with a larger average effect size in studies using relevant stimuli, $k = 26$, $n = 831$, $d = 1.595$, 95% CI [1.336, 1.854], compared with studies using less relevant stimuli, $k = 8$, $n = 218$, $d = 0.743$, 95% CI [0.319, 1.166].

Motivation to deceive had a statistically significant effect on the average effect size, $Q(2) = 6.439$, $p = .040$. However, there were only 6 studies in which participants were motivated with an incentive. While post hoc comparisons revealed no significant differences between studies in which participants were motivated with an incentive and the other two categories ($ps > .23$), they did reveal a smaller average effect size in studies in which it was reported that participants were motivated to lie as convincingly as possible compared with studies in which no extra motivation for participants was reported, $Q(1) = 6.036$, $p = .014$. This finding is in contrast with the results reported in CIT studies using ANS (Ben-Shakhar & Elaad, 2003; Meijer et al., 2014) and ERP measures (Ellwanger et al., 1996), and we therefore additionally examined the effect of motivation using only CIT studies. This analysis, in which we excluded the 2 CIT studies that reported having motivated participants with an incentive, revealed a still numerically smaller average effect size for CIT studies that reported motivational instructions, k = 10, $n = 274$, $d = 1.110$, 95% CI [0.684; 1.536], compared with studies in which no motivational instructions were reported, $k = 22$, $n = 723$, $d = 1.386$, 95% CI [1.084; 1.687], yet here the difference was not significant, $Q(1) = 1.073$, $p = .300$.

Dependency Analyses

Finally, to check for dependencies between the assessed categorical variables, we calculated Cramer's V as measure of the strength of the association between moderator variables. Cramer's V also serves as an effect size, with .10, .30, and .50 as thresholds

Table 3
Distribution (Absolute Numbers) and Dependencies (Cramer's V) of Moderator Variables

Variable	Paradigm					Motivation				Intertrial interval			
	CIT	aIAT	SLT	DoD	*V*	None	Instructions	Incentive	*V*	<4000 ms	4000–8000 ms	>8000 ms	*V*
Publication status													
Unpublished	0	0	8	0	.29*	8	0	0	.16	5	0	0	.19
Published	34	9	47	16		77	23	6		51	21	16	
RTs prim/sec measure													
No	20	0	8	13	.58***	27	8	6	.32**	17	9	10	.25
Yes	14	9	47	3		58	15	0		39	12	6	
Speed instructions													
No	2	0	7	2	.15	9	2	0	.09	5	5	1	.24
Yes	32	9	45	14		73	21	6		51	13	15	
Proportion lie/truth trials													
Equal proportion	0	0	37	10	.74***	42	2	3	.38**	11	16	11	.53***
More truth trials	34	9	11	5		36	21	2		43	5	5	
Stimulus relevance													
Low	8	5	46	12	.54***	60	8	3	.30**	21	15	15	.45***
High	26	4	9	4		25	15	3		35	6	1	
Motivation													
None	22	0	49	14						41	16	14	
Instructions	10	9	3	1	.45***					14	2	1	.21
Incentive	2	0	3	1						1	3	1	
Intertrial interval													
<4000 ms	25	3	24	4		41	14	1					
4000–8000 ms	5	0	15	1	.42***	16	2	3	.21				
>8000 ms	4	0	3	9			14	1		1			

Note. Instructions = Motivating instructions.
* $p < .05$. ** $p < .01$. *** $p < .001$.

for small, moderate, and large associations. As can be seen in Table 3, dependencies tended to be medium to large. Most importantly, we observed significant associations between motivation and paradigm (Cramer's $V = .45$, $p < .001$), motivation and stimulus proportion (Cramer's $V = .38$, $p < .01$), and paradigm and stimulus proportion (Cramer's $V = .74$, $p < .001$). The distribution of most variables was unequal for the different paradigms.

Countermeasure Meta-Analysis

From an applied perspective, it is important to determine whether the use of countermeasures allows participants to influence their results in RT deception paradigms. Therefore, we conducted an additional meta-analysis summarizing studies that fulfilled our inclusion criteria (except for criterion 8) but in which participants were instructed to employ countermeasures. Studies were coded as countermeasure studies when information about the test rationale or how to beat the test was given to participants. We included all studies that measured and reported RTs, independent of whether RTs were the primary or a secondary outcome measure. This also means that countermeasure instructions did not need to be specific for RTs. Examples of RT-specific countermeasure instructions are slowing down truth RTs (Seymour et al., 2000; Verschuere, Prati, & De Houwer, 2009). Examples of non RT-specific countermeasures are applying pressure to toes during truth trials (Mertens & Allen, 2008). As can be seen in Table 4, this search resulted in 17 independent studies, with 384 participants. Of those studies, 9 used the aIAT, 5 used the CIT, and 3 studies used the SLT. Twelve studies measured RTs as primary measures (9 aIATs, 2 SLTs, and 1 CITs), and 5 studies measured RTs as secondary measures (4 CITs, 1 SLT). Summarizing those studies resulted in a small and nonsignificant effect size, $d = 0.128$, 95% CI [−0.172; 0.429]. The Egger test did not reveal an indication for publication bias (value of the intercept = −0.186, $p = .941$). Cochran's Q revealed a large heterogeneity in the sample of studies, $Q(16) = 104.162$, $p < .001$ and I^2 indicated that about 85% of the observed variance between effect sizes was caused by systematic differences between studies. Because of the relatively small sample size, we refrained from conducting further moderator analyses.[7]

Discussion

Lie detection has enormous potential, but faces great challenges. Facing the limitations of traditional lie detection approaches and formulating a psychological theory of how and under which circumstances lying and truth telling differ may be a first step toward

[7] It would have been preferable to directly run a meta-analysis on the reduction of the RT deception effect through countermeasures in all studies that directly contrasted these two conditions (average d in control conditions minus average d in countermeasures conditions). This was unfortunately not possible because of too much variation in the way such control conditions were realized in the respective studies (from studies with no control conditions to studies using within- or between-subjects designs). To investigate whether this drastic reduction in effect size may be explained by characteristics of the particular studies in which countermeasures were studied (e.g., the specific paradigms that were used), we also calculated the average effect size of all available direct control conditions in the countermeasure studies. While numerically smaller, the average effect size, $d = 0.818$, 95% CI [0.586; 1.050] of this meta-analysis ($k = 11$, $n = 206$) was large and did not differ from the full sample, $d = 1.049$; 95% CI [0.930; 1.169].

Table 4
Summary of the Studies Included in the Countermeasure Meta-Analysis and the Average Effect Size Across Studies

Study	Paradigm	*n*	*d*	95% CI *LL*	95% CI *UL*
Agosta et al. (2011) - Experiment 1	aIAT	14	.514	−.043	1.072
Agosta et al. (2011) - Experiment 2	aIAT	20	.479	.016	.941
Agosta et al. (2011) - Experiment 3	aIAT	34	.904	.505	1.303
Agosta et al. (2011) - Experiment 4	aIAT	12	−.117	−.685	.451
Ganis et al. (2011)	CIT	12	−2.569	−3.743	−1.396
Hu, Chen, & Fu (2012) - Instruction group	SLT	16	1.069	.455	1.683
Hu, Chen, & Fu (2012) - Training group	SLT	16	.362	−.144	.868
Hu, Rosenfeld, & Bodenhausen (2012)	aIAT	16	−.240	−.737	.257
Hu, Rosenfeld, & Bodenhausen (2012)	aIAT	16	−.933	−1.520	−.346
Huntjens, Verschuere. & McNally (2012) - Simulators	CIT	23	.481	.049	.913
Mertens & Allen (2008) - CM Group 1	CIT	18	.579	.080	1.078
Mertens & Allen (2008) - CM Group 2	CIT	15	1.153	.500	1.806
Mertens & Allen (2008) - CM Group 3	CIT	15	.504	−.033	1.041
Uncapher et al. (2015)	SLT	24	.223	−.185	.631
Verschuere, Prati, & De Houwer (2009) - Experiment 1	aIAT	18	−.634	−1.140	−.127
Verschuere, Prati, & De Houwer (2009) - Experiment 2	aIAT	37	−.393	−.727	−.058
Verschuere, Prati, & De Houwer (2009) - Experiment 3	aIAT	42	−.214	−.520	.092
Average effect size (random effects model)		**348**	**.128**	**−.172**	**.429**

Note. n = number of participants; d = standardized paired difference; 95% CI = 95% Confidence Interval; *LL* = Lower Limit; *UL* = Upper Limit.

the development of evidence-based lie tests. An emerging theory of deception holds that lying is typically more cognitively demanding than truth telling. We included 114 studies with 3307 participants in our meta-analysis to test the key hypothesis of the cognitive perspective of deception that lying takes time.

Our meta-analysis revealed a significant and large standardized paired RT difference between truthful and deceptive responses. This result may seem surprising as two earlier meta-analyses (DePaulo et al., 2003; Zuckerman et al., 1981) came to the opposite conclusion. These other meta-analyses, however, investigated a number of different verbal and nonverbal behavioral deception indices during interviews. As pointed out by Verschuere et al. (2015), RT measurement conditions in interview situations are suboptimal. RTs require precise measurement, a design that allows participants to respond immediately after the stimulus presentation, and a sufficient number of observations. Accordingly, we restricted our analysis to studies that measured RT in such optimal conditions. Consequently, this resulted in none of the studies included in the two previous meta-analyses being included in the current meta-analysis. The RT deception effect may thus be restricted to computerized paradigms that fulfill minimal quality criteria for RT recordings. The effect does not necessarily generalize to face to face interview contexts.

A concern for the interpretation of the results of our meta-analysis is the presence of publication bias, with mostly positive results finding their way into academic journals. Therefore, the average RT deception effect in our meta-analysis is likely to overestimate the true effect. Nevertheless, there are some indications that the overestimation may not be substantial. After correcting for publication bias using the trim and fill method, our analysis still revealed a large and significant estimate for the RT deception effect (d = 1.049). The presence of a publication bias is also substantiated by the observation that effect sizes were larger in studies reporting RT as a primary measure compared with studies reporting RT as a secondary measure. It seems reasonable to assume that publication bias is especially present in the former, whereas the latter do not rely as much on the statistical significance of their RT results. Nevertheless, the average effect size in studies using RTs as secondary measure remains large (d = 1.064), and is in fact comparable with the estimate for the corrected effect size provided earlier. We did not find a decline of the effect size in studies with RT as the primary measure. We had expected this as such an effect has been observed in many research areas (Ioannidis, 2005). We can only speculate about this unexpected finding. As mentioned above, RT measures of deception have been faced with much skepticism for a long time. It is not until recently that more and more attention and trust has been placed in RTs as measures of deception. The increased use of RT deception measures is likely to have improved the quality of the measurements, paradigms, and analyses.

Despite the presence of a publication bias in the RT deception literature, the impact of this bias seems to be only modest, and correcting for this bias still revealed a large and significant RT deception effect. Even so, the observed publication bias calls for our attention. Publication bias is observed in many research areas, but it may have especially severe consequences in applied deception detection settings. Getting unbiased estimates of effect sizes and correct classification rates is essential to make informed decisions as to whether or not deception detection techniques should be implemented in applied contexts, as well as to determine the weight that should be given to their results (e.g., by police investigators, juries, or judges). Fortunately, attention to publication bias has grown (Ioannidis, 2006; Nosek, Spies, & Motyl, 2012). Study preregistration may help to prevent publication bias (Pashler & Wagenmakers, 2012) and the importance of reporting and publishing all well-conducted research, irrespective of whether it yielded statistically significant results, cannot be overstated.

What Moderated the RT Deception Effect?

The RT deception effect varied substantially between studies. We were able to identify two important moderators of the RT deception effect: The **paradigm** and **motivational instructions** to lie as convincingly as possible.

By providing RT differences and average effect sizes for each of the four included paradigms, our meta-analysis provides benchmarks for future research and power analyses. Comparisons between the different types of paradigms revealed a significantly smaller effect size for the aIAT, compared with the other three paradigms. This is remarkable considering the very large accuracy rates that are reported in several aIAT studies (above 90% in the review of Agosta & Sartori, 2013). One explanation for the smaller effect size of the aIAT compared with the other paradigms may be found in its specific design. Whereas the other three paradigms rely, in most cases, on designs in which the two contrasted conditions (truth telling and lying) are presented randomly intermixed, truthful and deceptive conditions in the aIAT are presented in two different blocks. As a consequence, the aIAT may be more vulnerable to participants' attempts to control their behavior, as they can try to apply the same strategy (e.g., speeding up their deceptive responses or slowing down their truthful responses) to an entire block, instead of having to flexibly switch between strategies. Studies have indeed shown that the aIAT is vulnerable to such strategic manipulations (Agosta et al., 2011; Hu et al., 2012; Verschuere et al., 2009). It is noteworthy to further mention that so far, most aIAT studies were conducted by the developers of the aIAT. Studies completed in the same lab may be more similar to each other than studies completed in different labs. It therefore remains important to attempt to replicate these results in different laboratories (for a recent direct independent replication see Kleinberg & Verschuere, in press).

Whereas the aIAT resulted in a significantly lower average RT effect size compared with the CIT, the SLT, and the DOD, we found no significant differences between the latter three paradigms. This finding may be surprising given that those paradigms vary in the type of stimuli that are used (e.g., words vs. sentences), the control conditions (e.g., different vs. same stimuli), and the presumed underlying theoretical mechanisms (e.g., the orienting response for the CIT, executive functioning for the DoD and the SLT). Nevertheless, these paradigms share the same contrast between lying and truth telling. Our meta-analysis also revealed a numerically lower effect size for the RT CIT (d = 1.297, 95% [1.060; 1.535]), compared with the average effect size for the most often used ANS concealed information measure, skin conductance (d = 1.55, 95% [1.41; 1.69] and d = 1.55, 95% [1.44; 1.66] in Ben-Shakhar & Elaad, 2003, and Meijer et al., 2014, respectively). However, this comparison should be treated with caution because the meta-analyses fundamentally differ with regard to their focus on within- or between-subjects designs. Whereas we included only within-subject comparisons, the previous CIT meta-analyses either combined between and within-subject designs or only included between-subjects designs. To compare skin conductance and RT measures, more research is needed in which these measures are simultaneously assessed, preferably under the most optimal conditions for each measure (e.g., Verschuere, Crombez, Degrootte, & Rosseel, 2010). Combining RTs and skin conductance in concealed information tests seems promising, as recent research suggests a dissociation between those measures: RTs seem to primarily tap into response inhibition processes, whereas skin conductance measures seem to tap more into orienting processes (Klein Selle et al., 2016; Suchotzki, Verschuere, et al., 2015).

When participants were motivated to lie as convincingly as possible or to try to beat the test (without being taught countermeasures), the average RT deception effect was smaller. This finding is at odds with the motivation impairment theory (DePaulo et al., 1988). The motivation impairment effect has initially been observed in research using verbal and nonverbal cues of deception, where a stronger motivation to lie paradoxically led to stronger differences between lying and truth telling (DePaulo et al., 2003), possibly because of greater pressure to lie successfully. The motivation impairment effect has also been observed in CIT studies using skin conductance measures (Ben-Shakhar & Elaad, 2003; Meijer et al., 2014), but not with RTs (Kleinberg & Verschuere, 2015) or event-related potentials (Ellwanger et al., 1996). There are at least two possible explanations for the detrimental role of motivation on the RT deception effect. First, instructions to beat the test may have led to more and better attempts of participants to employ countermeasures and fake a truthful test outcome (Ben-Shakhar, 2011). Second, motivation has been shown to improve executive functioning (Hajcak, Moser, Yeung, & Simons, 2005; Inzlicht & Schmeichel, 2012; Legault & Inzlicht, 2013; Muraven & Slessareva, 2003), and may thus facilitate a cognitively demanding task such as lying. Motivated liars may be more focused on their task goal (i.e., to lie effectively; Debey et al., 2012), pay more attention to the task at hand (i.e., lying), and invest more effort in self-control and the effective suppression of the prepotent truth response (Spence et al., 2001). Our finding also provides empirical support for the role of motivation in Walczyk et al.'s ADCAT and their prediction that the motivation to deceive determines the amount of cognitive resources that participants are willing to invest in lying (Walczyk et al., 2014). These two possible explanations for the effect of motivation can be experimentally differentiated. If the motivation effect is related to better executive functioning, motivated participants should be characterized by faster lie responses compared with unmotivated participants. If the motivation effect is related to more successful faking attempts, motivated participants should show RT slowing for truth responses.

We found a significant effect of the proportion of truth versus lie trials, yet surprisingly studies using an equal proportion of truth and lie responses tended to result in a larger average effect size compared with studies using more truth trials. A closer inspection of the data revealed that the proportion factor was confounded with the motivation factor: Studies with more truth trials were more likely to be studies in which participants were instructed to lie as convincingly as possible. The use of motivational instructions resulted in a smaller effect size and this may have counteracted the expected proportion effect.

Because of the small number of studies, no moderator analyses could be conducted on the role of stimulus attention and population. The length of the intertrial interval, stimulus relevance, and the absolute number of trials did not moderate the RT deception effect. It is important to note that the *absence of an effect* of these variables cannot be taken as *evidence for no effect*, and some of those moderator effects may become significant with increased power. For instance, our data were not entirely conclusive on whether instructing participants to respond as fast as possible results in a larger average effect size or not ($p = .081$), and more primary research also manipulating such instructions within one study is necessary.

Is Lying Always More Demanding Than Truth Telling?

A remarkable finding in our main meta-analysis is that none of the significant moderators led to a reversed RT deception effect. Even when a moderator significantly lowered the effect sizes, effects remained positive and large. Moreover, not a single study in our main meta-analysis revealed a significant reversed RT lie effect (i.e., significantly longer RT for truth telling compared with lying, with the 95% CI not including zero, see Table 1).[8] This pattern of results is interesting in light of the current debate about whether lying is always more effortful than truth telling or whether—and if so under which circumstances—truth telling may be more effortful than lying (Levine, 2014; McCornack et al., 2014; Verschuere & Shalvi, 2014). Our findings revealed a surprising stability of the cognitive cost of deception, even in some of the situations that have been proposed to modulate or reverse this cost (e.g., practiced deception, relevant information, and motivated liars; DePaulo et al., 2003; McCornack et al., 2014; Walczyk et al., 2014). Evidently, laboratory research is restricted in the degree to which factors such as motivation or the relevance of the information being lied about can be manipulated. We cannot exclude the possibility that more extreme levels of those factors may have a stronger influence on the cost associated with deception.

The results of our meta-analysis also contribute to the question whether the truth or the lie constitutes the default mode in human communication. According to the cognitive view on deception, the truth is typically activated first, and lying requires overcoming of the automatic truth response (e.g., Christ et al., 2009; Spence et al., 2001; Verschuere et al., 2011). In contrast, researchers from social psychology and behavioral economics have pointed out that in tempting situations, lying may be the more automatic response, and it is truth telling that may require cognitive effort and deliberation (e.g., Ariely, 2012; Bereby-Meyer & Shalvi, 2015; Shalvi, Eldar, & Bereby-Meyer, 2012). Indeed, given the opportunity to gain a monetary reward, it has been observed that people deceive more when deliberation is limited (e.g., by fatigue or time pressure; Kouchaki & Smith, 2014; Shalvi et al., 2012). Crucially, the cognitive studies and the social psychology/behavioral economics studies differ in at least two important aspects. First, whereas the cognitive studies examine the cognitive effort associated with truthful and deceptive communication, the social psychology and behavioral economics studies look whether and under which circumstances people deceive. This difference could be related to Walczyk et al.'s (2014) ADCAT, and one may argue that they investigate different stages of the deception process: The decision to deceive versus the cognitive effort associated with the subsequent construction and action of the deceitful communication. Second, the two lines of research differ in the incentives used (Verschuere & Shalvi, 2014). Studies from cognitive psychology often use motivational instructions but often do not financially reward successful deception. In contrast, most studies from social psychology and behavioral economics create tempting situations, in which the decision to deceive is financially rewarded. Keeping this difference in mind, combining findings from both types of research could indicate that motivation influences the deception process differently at the different stages. The presence of a high motivation to deceive, for instance in situations in which lying is easy and financially beneficial, may especially facilitate the decision to lie (Kouchaki & Smith, 2014; Shalvi et al., 2012). Once the decision to lie or tell the truth is made, however, the results of our meta-analysis suggest that a higher motivation to lie does not eliminate the cognitive effort that comes with the construction of the lie and the deceptive action. Combining the two perspectives and more systematically investigating all deception stages could be a fruitful avenue for future research. For instance, more studies in which participants can freely choose whether to lie or tell the truth should be conducted (Panasiti et al., 2014; Pfister et al., 2014; Wu et al., 2009), and the impact of this decision process on the RT cost of deception should be studied (Walczyk, Roper, Seemann, & Humphrey, 2003). Another interesting question is whether there is a relatively lower cognitive cost of deceptive compared with truthful responding for people with a higher propensity to cheat in tempting situations (Abe & Greene, 2014; Greene & Paxton, 2009; Hu, Pornpattananangkul, & Nusslock, 2015; Tabatabaeian, Dale, & Duran, 2015).

What Underlies the Enhanced Cognitive Cost of Deception?

The finding that lying on average takes longer than truth telling supports the idea that lying has a cognitive cost. Three executive functions have been proposed to underlie this cognitive cost: Working memory, response inhibition, and task switching. Each of these executive functions may contribute to the RT deception effect. In support of the working memory hypothesis, it has been shown that during lying, the truth needs to be activated first, elevating RTs for lying (Debey, De Houwer, & Verschuere, 2014). Research has also shown that the concurrent activation of the truth and the lie response results in a conflict between the two responses (e.g., Dong, Hu, Lu, & Wu, 2010; Johnson, Henkell, Simon, & Zhu, 2008; Vidal, Hasbroucq, Grapperon, & Bonnet, 2000). Tracking participants' arm movements while moving a Nintendo Wii Remote to truthful or deceptive response options, Duran et al. (2010) observed longer RTs for lying compared with truth telling, due to a stronger initial deviation of the trajectories toward the truthful response. Although the conflict inducing nature of deception has been supported, research is inconclusive on whether active response inhibition is required to overcome the prepotent truth response (Debey, Ridderinkhof, De Houwer, De Schryver, & Verschuere, 2015; Duran, Dale, & McNamara, 2010; Hadar, Makris, & Yarrow, 2012; Suchotzki, Crombez, Debey, Van Oorsouw, & Verschuere, 2015; Vartanian, Kwantes, & Mandel, 2012).

[8] Here we do not refer to studies in which countermeasure were employed. Countermeasures are strategies employed by subjects to strategically alter lie test outcomes, and thereby mask the cognitive dynamics of lying. Note that employing countermeasures is not equivalent with practiced deceptive responses, with the latter being included in our main meta-analysis.

Finally, switching between truth telling and lying, like switching between different tasks (Kiesel et al., 2010; Monsell, 2003; Vandierendonck, Liefooghe, & Verbruggen, 2010) could elevate RTs—though it has been shown that the RT difference between lying and truth telling is not merely a switch cost (Debey, Liefooghe, De Houwer, & Verschuere, 2015).

It will be important to disentangle the role of these three executive functions during deception. Correlating independent measures of working memory, response inhibition, and task switching with RT deception effects could reveal information about their relative contributions (Visu-Petra, Miclea, & Visu-Petra, 2012). Selectively interfering with each of these functions (e.g., with parallel working memory tasks or by depleting inhibition resources) may increase RT deception effects and provide more direct support for the causal role of the respective functions in deception (Ambach, Stark, Peper, & Vaitl, 2008b; Ambach, Stark, & Vaitl, 2011; Debey et al., 2012; Verschuere et al., 2012; Visu-Petra, Varga, Miclea, & Visu-Petra, 2013). Also, future research should develop and test predictions that are specific for each of the executive functions. For instance, successful inhibition of the truth response should lead to a leveling-off of the increase of the RT deception effect with longer RTs (i.e., the delta plot method; Debey, Ridderinkhof, et al., 2015; Ridderinkhof, van den Wildenberg, Wijnen, & Burle, 2004).

Is It Purely Cognitive?

Our results are in line with a cognitive view of deception. Nevertheless, the RT difference between lying and truth telling is not necessarily purely driven by cognitive mechanisms. Even when matching the valence and arousal of stimuli in the truth and lie conditions as closely as possible, like many studies in the current meta-analysis did, deception may be intrinsically more emotional and arousing than truth telling. Indeed, in a series of experiments using the DoD paradigm, Furedy et al. found stable differences in autonomic arousal between lying and truth telling toward well matched neutral questions (Furedy et al., 1988; Furedy, Gigliotti, & Ben-Shakhar, 1994; Furedy, Posner, & Vincent, 1991), and these differences seemed to be unaffected by cognitive load (Vincent & Furedy, 1992). In the so-called emotional Stroop task, emotional words have been found to prolong RTs for color naming responses compared with neutral words (Algom, Chajut, & Lev, 2004). Correspondingly, stronger emotional arousal for deception compared with truth telling may also contribute to RT deception effects. It is therefore important to examine the role of emotion and its possible interaction with cognition in deception more closely (Dolcos, Iordan, & Dolcos, 2011; Pessoa, 2008). Such research may, for instance, compare lying about neutral and emotional stimuli to explore effects of stimulus valence on the RT for truth telling and lying (Lee, Lee, Raine, & Chan, 2010; Proverbio, Vanutelli, & Adorni, 2013). Another possibility would be to manipulate the extrinsic emotional valence of lying and truth telling, for instance by associating one behavior with positive consequences (e.g., financial reward) and the other one with punishment (e.g., aversive noises or shocks, see, e.g., Tomash & Reed, 2015). One may also try to assess the intrinsic emotional valence of lying with measures that are sensitive to valence and arousal, such as the startle eyeblink (Verschuere, Crombez, Koster, Van Bockstaele, & De Clercq, 2007) or affective priming measures (Aarts, De Houwer, & Pourtois, 2012; Fazio, 2001).

Practical Implications

Can RTs be used for lie detection? The large overall average effect indicates that RTs may have potential in applied settings. However, the DoD and the SLT have been nearly exclusively used for basic deception research. Their use in applied settings has been questioned (Furedy et al., 1988), because participants cannot be expected to comply with the truth telling and lying instruction in such contexts. Thus, DoD and the SLT require modification to be used in the applied context. A first attempt to introduce the SLT in a forensic context was made by Spence, Kaylor-Hughes, Brook, Lankappa, and Wilkinson (2008) who used “denying” versus “admitting” instructions rather than lying versus truth telling. Thus, the examinee was asked in one block to admit to the allegations and in another block to deny the allegations. Spence at al. expected the condition that represented the lie to result in longer RTs and more neuronal activation in brain regions associated with lying. In this single case study, both the RT and the neural deception effect pointed in the same direction (toward the innocence of the suspect), yet this result could not be verified because ground truth was not available. Unfortunately, further research on the diagnostic value of the modified SLT is not yet available.

The CIT and the aIAT could more readily be used in applied settings. The RT CIT can be applied to detect recognition of concealed (e.g., crime) information. The aIAT can be applied to find out which of two contrasting events (e.g., a crime vs. an alibi) is true. It should be mentioned here that the present meta-analysis focused exclusively on within-subject comparisons. This choice was a deliberate one, because we think that a within-subject comparison of lie and truth is very useful for RT-based lie detection given the great interindividual variability in RTs. There are many cases in which such within-subject results can provide valuable information in applied situations (e.g., contrasting a crime and an alibi story to evaluate which story a suspect is lying about). Still, the diagnostic efficiency of a RT-based lie detection test will require a comprehensive estimation of both its sensitivity and its specificity.

A comparison of sensitivity and specificity for the RT CIT and the CIT using skin conductance, event-related potentials, and functional magnetic resonance imaging (fMRI) is available in a recent review of Meijer et al. (2016). In this review, the area (a) under the receiver operating curve (ROC) was used as measure of diagnostic accuracy. The advantage of this index is its independence of specific cut-off points, which may differ across studies (see also the recommendation of the National Research Council, 2003). Instead, a depicts the classification accuracy across all possible cut-off points, with values ranging from .50 (chance classification) to 1 (perfect classification). The review by Meijer et al. found a weighted average of $a = .82$ for the RT CIT ($n = 981$), of $a = .85$ for the skin conductance CIT ($n = 3863$), of $a = .88$ for the event related potential CIT ($n = 646$), and of $a = .94$ for the fMRI-based CIT ($n = 134$). Except for the skin conductance CIT, estimates were based on a relatively small number of participants. Also, all included studies were conducted in the laboratory, leading to a restricted external validity. More research should aim to increase the external validity and the precision of the estimates to determine whether any differences reflect genuine differences between measures.

Given that RTs are in principle under voluntary control, it is important to test their vulnerability to faking attempts. To this end, we conducted an additional meta-analysis summarizing research on the impact of countermeasures on the RT deception effect. The average effect size of the RT deception effect in countermeasures studies was small and nonsignificant, indicating that RT deception measures are highly vulnerable to countermeasures. An important qualification is that 9 of the 17 studies used the aIAT which produced smaller effects in the main meta-analysis and in which it may be easier to fake an innocent outcome than in the other RT paradigms (Agosta et al., 2011; Hu et al., 2012; Verschuere et al., 2009). Interestingly, the main meta-analysis showed that although motivation reduces the RT deception effect, it is not sufficient to wash out/mask the effect: The RT deception effect remained large even when participants were motivated to lie as convincingly as possible. Taken together, these findings suggest that motivation is not enough to fake a truthful test outcome. Successful faking seems to require knowledge about the test principles and/or specific instructions on how to manipulate responses. Two important avenues for future research are the development of algorithms that may enable the detection of countermeasure use (Agosta et al., 2011; Cvencek, Greenwald, Brown, Gray, & Snowden, 2010; Rosenfeld et al., 2008) and the identification of design characteristics that may hamper countermeasure use (Rosenfeld et al., 2008; Verschuere & Meijer, 2014).

A disappointing finding with regard to the application of RT-based measures of deception was that there were only four studies in clinical or forensic populations. Of those, two studies were conducted with a clinical sample (patients with essential tremor and patients with Parkinson's disease; Mameli et al., 2013), and two studies were conducted with forensic-clinical samples. The RT deception effect and the neural mechanisms during deception were studied in a sample of youth offenders with an additional diagnosis of antisocial personality disorder (Jiang et al., 2013), and in a sample of participants with schizophrenia, of which two thirds had previous contact with the police (Kaylor-Hughes et al., 2011). Based on the idea that successful lying crucially relies on executive functions (Christ et al., 2009), which can be impaired in clinical and forensic populations (Miyake et al., 2000), one may expect larger deception RT effects in clinical and forensic samples. Knowledge on how effect sizes vary in different populations is crucial for an application of RT tests and this type of research requires a major effort in the future.

Limitations

The first limitation concerns our moderator analyses. In meta-analyses, levels of moderators are not randomly assigned to studies and consequently moderator effects may be confounded with other variables. The results revealed that in our moderator analyses, the significant moderator motivation was correlated with paradigm and stimulus proportion. More primary research manipulating these variables within controlled experimental designs is needed before strong conclusions can be drawn about how these variables may have influenced the motivation effect and each other.

A second limitation concerns the operationalization of deception in the current meta-analysis as well as in the deception field in general. According to many definitions of lying (e.g., Burgoon & Buller, 1994; Krauss, 1981; Vrij, 2008; Zuckerman et al., 1981) deception involves making someone else believe, without forewarning, what the deceiver considers to be untrue. Yet, in many deception studies, participants are instructed by the experimenter when to lie and when to tell the truth. Also, in many studies, participants aim to deceive a computer, lacking the social component that many definitions stress. More generally, there is a need to more closely investigate the differences between deception in a laboratory and in real-life (Sip, Roepstorff, McGregor, & Frith, 2008).

A third limitation is that we could not test a factor that was proposed to be crucial in RT deception measures: Response conflict. Whether or not a paradigm manipulates the conflict between the truthful response evoked by the stimulus and the required deceptive response during lie trials has been proposed to explain inconsistent results with different RT deception paradigms (Suchotzki et al., 2013; Verschuere & De Houwer, 2011). We considered coding response conflict as a possible moderating factor, but realized that by including only studies in which truth telling and lying is manipulated within subjects and in which RTs are measured, our sample nearly exclusively consisted of studies in which response competition was inherent to the task. For instance, all but one of the CIT studies included in our sample used (variations of) the 3-stimulus protocol. Here, response competition occurs between truthfully denying knowledge of irrelevant items and truthfully acknowledging recognition of target items, and deceptively denying knowledge of the critical items. All aIAT studies included in our sample rely on the response competition occurring when autobiographically true statements have to be combined with generally false statements (and vice versa). All SLT and DoD studies included in our meta-analysis relied on the competition between the truthful response and the required deceptive response during lie trials. It seems worth pointing out that we did not deliberately exclude studies without response conflict, but that this selection is partly a natural consequence of concentrating on paradigms that manipulate and measure deception directly, as such tasks always involve the conflict between truth telling and lying.

A fourth limitation concerns the conclusions that can be drawn from the nonsignificant results of our moderator analysis. As noted earlier, absence of evidence for an effect cannot be taken as evidence for no effect. The large confidence intervals indicated that more research is needed to know whether some moderator effects will reach significance with increased power. Also, the average sample size in the studies included in this meta-analysis was modest ($n = 29$). Larger sample sizes may further increase the precision of effect estimates. Online research seems an efficient and promising way to conduct well-powered RT studies on deception (Kleinberg & Verschuere, 2015).

Conclusions

The current meta-analysis revealed a large and significant overall RT deception effect in computerized deception paradigms, even after controlling for publication bias. RT thus appears to be an effective measure for the study of deception. The RT paradigms included in our meta-analysis provide strict experimental control for the crucial lie-truth contrast. The obtained RT deception effect fits with the contemporary and empirically supported idea that lying is cognitively more demanding than truth telling. Whether or not RT-based lie detection is applicable in forensic settings remains to be determined in empirical research, carefully mapping its boundary conditions, external validity, and exploring solutions regarding its vulnerability to countermeasures. We hope that deception researchers will pick up the challenge. After all, it is only the scientific enterprise that can improve the status quo and differentiate valid from invalid lie detection methods.

References

References marked with an asterisk indicate studies included in the main meta-analysis. References marked with a hashtag indicate studies included in the countermeasure meta-analysis.

Aarts, K., De Houwer, J., & Pourtois, G. (2012). Evidence for the automatic evaluation of self-generated actions. *Cognition, 124,* 117–127. http://dx.doi.org/10.1016/j.cognition.2012.05.009

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

Abe, N., & Greene, J. D. (2014). Response to anticipated reward in the nucleus accumbens predicts behavior in an independent test of honesty. *The Journal of Neuroscience, 34,* 10564–10572. http://dx.doi.org/10.1523/JNEUROSCI.0217-14.2014

*Abe, N., Okuda, J., Suzuki, M., Sasaki, H., Matsuda, T., Mori, E., . . . Fujii, T. (2008). Neural correlates of true memory, false memory, and deception. *Cerebral Cortex, 18,* 2811–2819. http://dx.doi.org/10.1093/cercor/bhn037

Abe, N., Suzuki, M., Tsukiura, T., Mori, E., Yamaguchi, K., Itoh, M., & Fujii, T. (2006). Dissociable roles of prefrontal and anterior cingulate cortices in deception. *Cerebral Cortex, 16,* 192–199. http://dx.doi.org/10.1093/cercor/bhi097

*#Agosta, S., Ghirardi, V., Zogmaister, C., Castiello, U., & Sartori, G. (2011). Detecting Fakers of the autobiographical IAT. *Applied Cognitive Psychology, 25,* 299–306. http://dx.doi.org/10.1002/acp.1691

Agosta, S., & Sartori, G. (2013). The autobiographical IAT: A review. *Frontiers in Psychology, 4,* 519. http://dx.doi.org/10.3389/fpsyg.2013.00519

Algom, D., Chajut, E., & Lev, S. (2004). A rational look at the emotional stroop phenomenon: A generic slowdown, not a stroop effect. *Journal of Experimental Psychology: General, 133,* 323–338. http://dx.doi.org/10.1037/0096-3445.133.3.323

*Allen, J. J., Iacono, W. G., & Danielson, K. D. (1992). The identification of concealed memories using the event-related potential and implicit behavioral measures: A methodology for prediction in the face of individual differences. *Psychophysiology, 29,* 504–522. http://dx.doi.org/10.1111/j.1469-8986.1992.tb02024.x

Ambach, W., Stark, R., Peper, M., & Vaitl, D. (2008a). Separating deceptive and orienting components in a Concealed Information Test. *International Journal of Psychophysiology, 70,* 95–104. http://dx.doi.org/10.1016/j.ijpsycho.2008.07.002

Ambach, W., Stark, R., Peper, M., & Vaitl, D. (2008b). An interfering Go/No-go task does not affect accuracy in a Concealed Information Test. *International Journal of Psychophysiology, 68,* 6–16. http://dx.doi.org/10.1016/j.ijpsycho.2007.11.004

Ambach, W., Stark, R., & Vaitl, D. (2011). An interfering n-back task facilitates the detection of concealed information with EDA but impedes it with cardiopulmonary physiology. *International Journal of Psychophysiology, 80,* 217–226. http://dx.doi.org/10.1016/j.ijpsycho.2011.03.010

Ariely, D. (2012). *The (honest) truth about dishonesty: How we lie to everyone—Especially ourselves*. New York, NY: Harper.

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study. *Psychological Bulletin, 133,* 1–24. http://dx.doi.org/10.1037/0033-2909.133.1.1

Ben-Shakhar, G. (1977). A further study of the dichotomization theory in detection of information. *Psychophysiology, 14,* 408–403. http://dx.doi.org/10.1111/j.1469-8986.1977.tb02974.x

Ben-Shakhar, G. (2002). A critical review of the Control Questions Test (CQT). In M. Kleiner (Ed.), *Handbook of polygraph testing* (pp. 103–126). Waltham, MA: Academic Press.

Ben-Shakhar, G. (2011). Countermeasures. In B. Verschuere, G. Ben Shakhar, & E. Meijer (Eds.), *Memory detection: Theory and application of the Concealed Information Test* (pp. 200–214). Cambridge, UK: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511975196.012

Ben-Shakhar, G., & Elaad, E. (2003). The validity of psychophysiological detection of information with the Guilty Knowledge Test: A meta-analytic review. *Journal of Applied Psychology, 88,* 131–151. http://dx.doi.org/10.1037/0021-9010.88.1.131

Ben-Shakhar, G., Lieblich, I., & Kugelmass, S. (1982). Interactive effects of stimulus probability and significance on the skin conductance response. *Psychophysiology, 19,* 112–114. http://dx.doi.org/10.1111/j.1469-8986.1982.tb02610.x

Bereby-Meyer, Y., & Shalvi, S. (2015). Deliberate honesty. *Current Opinion in Psychology, 6,* 195–198. http://dx.doi.org/10.1016/j.copsyc.2015.09.004

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2011). *Introduction to meta-analysis*. Chichester, UK: Wiley.

Burgoon, J. K., & Buller, D. B. (1994). Interpersonal deception: III. Effects of deceit on perceived communication and nonverbal behavior dynamics. *Journal of Nonverbal Behavior, 18,* 155–184. http://dx.doi.org/10.1007/BF02170076

Burgoon, J. K., & Floyd, K. (2000). Testing for the motivation impairment effect during deceptive and truthful interaction [includes Communication Reports]. *Western Journal of Communication, 64,* 243–267. http://dx.doi.org/10.1080/10570310009374675

Carmel, D., Dayan, E., Naveh, A., Raveh, O., & Ben-Shakhar, G. (2003). Estimating the validity of the guilty knowledge test from simulated experiments: The external validity of mock crime studies. *Journal of Experimental Psychology: Applied, 9,* 261–269. http://dx.doi.org/10.1037/1076-898X.9.4.261

*Carrión, R. E., Keenan, J. P., & Sebanz, N. (2010). A truth that's told with bad intent: An ERP study of deception. *Cognition, 114,* 105–110. http://dx.doi.org/10.1016/j.cognition.2009.05.014

Caso, L., Gnisci, A., Vrij, A., & Mann, S. (2005). Processes underlying deception: An empirical analysis of truth and lies when manipulating the stakes. *Journal of Investigative Psychology and Offender Profiling, 2,* 195–202. http://dx.doi.org/10.1002/jip.32

Chan, M. E., & Arvey, R. D. (2012). Meta-Analysis and the development of knowledge. *Perspectives on Psychological Science, 7,* 79–92. http://dx.doi.org/10.1177/1745691611429355

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-analyses. *Cerebral Cortex, 19,* 1557–1566. http://dx.doi.org/10.1093/cercor/bhn189

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences*. Hillsdale, NJ: Erlbaum.

Crombez, G., Van Ryckeghem, D. M., Eccleston, C., & Van Damme, S. (2013). Attentional bias to pain-related information: A meta-analysis. *Pain, 154,* 497–510. http://dx.doi.org/10.1016/j.pain.2012.11.013

Cumming, G., & Finch, S. (2005). Inference by eye: Confidence intervals and how to read pictures of data. *American Psychologist, 60,* 170–180. http://dx.doi.org/10.1037/0003-066X.60.2.170

*Cutmore, T. R. H., Djakovic, T., Kebbell, M. R., & Shum, D. H. K. (2009). An object cue is more effective than a word in ERP-based detection of deception. *International Journal of Psychophysiology, 71,* 185–192. http://dx.doi.org/10.1016/j.ijpsycho.2008.08.003

Cvencek, D., Greenwald, A. G., Brown, A. S., Gray, N. S., & Snowden, R. J. (2010). Faking of the implicit association test is statistically detectable and partly correctable. *Basic and Applied Social Psychology, 32,* 302–314. http://dx.doi.org/10.1080/01973533.2010.519236

Davatzikos, C., Ruparel, K., Fan, Y., Shen, D. G., Acharyya, M., Loughead, J. W., . . . Langleben, D. D. (2005). Classifying spatial patterns of brain activity with machine learning methods: Application to lie detection. *NeuroImage, 28,* 663–668. http://dx.doi.org/10.1016/j.neuroimage.2005.08.009

*Debey, E. (2014). *Behavioral studies on the role of cognitive control in deception* (Unpublished doctoral dissertation). Ghent University, Ghent, Belgium.

*Debey, E., De Houwer, J., & Verschuere, B. (2014). Lying relies on the truth. *Cognition, 132,* 324–334. http://dx.doi.org/10.1016/j.cognition.2014.04.009

*Debey, E., Liefooghe, B., De Houwer, J., & Verschuere, B. (2015). Lie, truth, lie: The role of task switching in a deception context. *Psychological Research, 79,* 478–488.

Debey, E., Ridderinkhof, R. K., De Houwer, J., De Schryver, M., & Verschuere, B. (2015). Suppressing the truth as a mechanism of deception: Delta plots reveal the role of response inhibition in lying. *Consciousness and Cognition, 37,* 148–159. http://dx.doi.org/10.1016/j.concog.2015.09.005

*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. http://dx.doi.org/10.1016/j.actpsy.2012.03.004

DePaulo, B. M., Kirkendol, S. E., Tang, J., & O'Brien, T. P. (1988). The motivational impairment effect in the communication of deception: Replications and extensions. *Journal of Nonverbal Behavior, 12,* 177–202. http://dx.doi.org/10.1007/BF00987487

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

Dolcos, F., Iordan, A. D., & Dolcos, S. (2011). Neural correlates of emotion-cognition interactions: A review of evidence from brain imaging investigations. *Journal of Cognitive Psychology, 23,* 669–694. http://dx.doi.org/10.1080/20445911.2011.594433

Donders, F. C. (1969). On the speed of mental processes. *Acta Psychologica, 30,* 412–431. http://dx.doi.org/10.1016/0001-6918(69)90065-1

Dong, G., Hu, Y., Lu, Q., & Wu, H. (2010). The presentation order of cue and target matters in deception study. *Behavioral and Brain Functions, 6,* 63. http://dx.doi.org/10.1186/1744-9081-6-63

Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods, 1,* 170–177. http://dx.doi.org/10.1037/1082-989X.1.2.170

*Duran, N. D., Dale, R., & McNamara, D. S. (2010). The action dynamics of overcoming the truth. *Psychonomic Bulletin & Review, 17,* 486–491. http://dx.doi.org/10.3758/PBR.17.4.486

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56,* 455–463. http://dx.doi.org/10.1111/j.0006-341X.2000.00455.x

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ, 315,* 629–634. http://dx.doi.org/10.1136/bmj.315.7109.629

Ekman, P. (1985). *Telling lies: Clues to deceit in the marketplace, marriage, and politics*. New York, NY: Norton.

Ellwanger, J., Rosenfeld, J. P., Sweet, J. J., & Bhatt, M. (1996). Detecting simulated amnesia for autobiographical and recently learned information using the P300 event-related potential. *International Journal of Psychophysiology, 23*(1–2), 9–23. http://dx.doi.org/10.1016/0167-8760(96)00035-9

Engelhard, I. M., Merckelbach, H., & van den Hout, M. A. (2003). The Guilty Knowledge Test and the modified Stroop task in detection of deception: An exploratory study. *Psychological Reports, 92,* 683–691. http://dx.doi.org/10.2466/pr0.2003.92.2.683

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. Ross (Ed.), *The psychology of learning and motivation* (pp. 145–199). New York, NY: Academic Press.

Eriksson, A., & Lacerda, F. (2007). Charlatanry in forensic speech science: A problem to be taken seriously. *International Journal of Speech Language and the Law, 14,* 169–193. http://dx.doi.org/10.1558/ijsll.v14i2.169

Estes, W. (1997). On the communication of information by displays of standard errors and confidence intervals. *Psychonomic Bulletin & Review, 4,* 330–341. http://dx.doi.org/10.3758/BF03210790

Fanelli, D. (2012). Negative results are disappearing from most disciplines and countries. *Scientometrics, 90,* 891–904. http://dx.doi.org/10.1007/s11192-011-0494-7

*Farrow, T. F. D., Hopwood, M.-C., Parks, R. W., Hunter, M. D., & Spence, S. A. (2010). Evidence of mnemonic ability selectively affecting truthful and deceptive response dynamics. *The American Journal of Psychology, 123,* 447–453. http://dx.doi.org/10.5406/amerjpsyc.123.4.0447

*Farrow, T. F. D., Reilly, R., Rahman, T. A., Herford, A. E., Woodruff, P. W. R., & Spence, S. A. (2003). Sex and personality traits influence the difference between time taken to tell the truth or lie. *Perceptual and Motor Skills, 97,* 451–460. http://dx.doi.org/10.2466/pms.2003.97.2.451

Farwell, L. A., & Donchin, E. (1991). The truth will out: Interrogative polygraphy ("lie detection") with event-related brain potentials. *Psychophysiology, 28,* 531–547. http://dx.doi.org/10.1111/j.1469-8986.1991.tb01990.x

Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition and Emotion, 15,* 115–141. http://dx.doi.org/10.1080/02699930125908

*Fecteau, S., Boggio, P., Fregni, F., & Pascual-Leone, A. (2013). Modulation of untruthful responses with non-invasive brain stimulation. *Frontiers in Psychiatry, 3,* 97. http://dx.doi.org/10.3389/fpsyt.2012.00097

*Fullam, R. S., McKie, S., & Dolan, M. C. (2009). Psychopathic traits and deception: Functional magnetic resonance imaging study. *The British Journal of Psychiatry, 194,* 229–235. http://dx.doi.org/10.1192/bjp.bp.108.053199

Furedy, J. J., Davis, C., & Gurevich, M. (1988). Differentiation of deception as a psychological process: A psychophysiological approach. *Psychophysiology, 25,* 683–688. http://dx.doi.org/10.1111/j.1469-8986.1988.tb01908.x

Furedy, J. J., Gigliotti, F., & Ben-Shakhar, G. (1994). Electrodermal differentiation of deception: The effect of choice versus no choice of deceptive items. *International Journal of Psychophysiology, 18,* 13–22. http://dx.doi.org/10.1016/0167-8760(84)90011-4

Furedy, J. J., Posner, R. T., & Vincent, A. (1991). Electrodermal differentiation of deception: Perceived accuracy and perceived memorial content manipulations. *International Journal of Psychophysiology, 11,* 91–97. http://dx.doi.org/10.1016/0167-8760(91)90376-9

Gamer, M. (2011). Detection of deception and concealed information using neuroimaging techniques. In B. Verschuere, G. Ben-Shakhar, & E. Meijer (Eds.), *Memory detection: Theory and application of the Concealed Information Test* (pp. 90–113). Cambrige, UK: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511975196.006

*Gamer, M., Bauermann, T., Stoeter, P., & Vossel, G. (2007). Covariations among fMRI, skin conductance, and behavioral data during processing of concealed information. *Human Brain Mapping, 28,* 1287–1301. http://dx.doi.org/10.1002/hbm.20343

*Gamer, M., & Berti, S. (2010). Task relevance and recognition of concealed information have different influences on electrodermal activity and event-related brain potentials. *Psychophysiology, 47,* 355–364. http://dx.doi.org/10.1111/j.1469-8986.2009.00933.x

*Gamer, M., & Berti, S. (2012). P300 amplitudes in the concealed information test are less affected by depth of processing than electrodermal responses. *Frontiers in Human Neuroscience, 6,* 308. http://dx.doi.org/10.3389/fnhum.2012.00308

Gamer, M., Kosiol, D., & Vossel, G. (2010). Strength of memory encoding affects physiological responses in the Guilty Actions Test. *Biological Psychology, 83,* 101–107. http://dx.doi.org/10.1016/j.biopsycho.2009.11.005

Ganis, G., & Keenan, J. P. (2009). The cognitive neuroscience of deception. *Social Neuroscience, 4,* 465–472. http://dx.doi.org/10.1080/17470910802507660

*Ganis, G., Morris, R. R., & Kosslyn, S. M. (2009). Neural processes underlying self- and other-related lies: An individual difference approach using fMRI. *Social Neuroscience, 4,* 539–553. http://dx.doi.org/10.1080/17470910801928271

*#Ganis, G., Rosenfeld, J. P., Meixner, J., Kievit, R. A., & Schendan, H. E. (2011). Lying in the scanner: Covert countermeasures disrupt deception detection by functional magnetic resonance imaging. *NeuroImage, 55,* 312–319. http://dx.doi.org/10.1016/j.neuroimage.2010.11.025

Ganis, G., & Schendan, H. E. (2013). Concealed semantic and episodic autobiographical memory electrified. *Frontiers in Human Neuroscience, 6,* 354. http://dx.doi.org/10.3389/fnhum.2012.00354

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. http://dx.doi.org/10.3200/MONO.132.3.197-214

Greene, J. D., & Paxton, J. M. (2009). Patterns of neural activity associated with honest and dishonest moral decisions. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 106,* 12506–12511. http://dx.doi.org/10.1073/pnas.0900152106

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85,* 197–216. http://dx.doi.org/10.1037/0022-3514.85.2.197

Gronau, N., Ben-Shakhar, G., & Cohen, A. (2005). Behavioral and physiological measures in the detection of concealed information. *Journal of Applied Psychology, 90,* 147–158. http://dx.doi.org/10.1037/0021-9010.90.1.147

Hadar, A. A., Makris, S., & Yarrow, K. (2012). The truth-telling motor cortex: Response competition in M1 discloses deceptive behaviour. *Biological Psychology, 89,* 495–502. http://dx.doi.org/10.1016/j.biopsycho.2011.12.019

Hajcak, G., Moser, J. S., Yeung, N., & Simons, R. F. (2005). On the ERN and the significance of errors. *Psychophysiology, 42,* 151–160. http://dx.doi.org/10.1111/j.1469-8986.2005.00270.x

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

Henke, F. G., & Eddy, M. W. (1909). Mental diagnosis by the association reaction method. *Psychological Review, 16,* 399–409. http://dx.doi.org/10.1037/h0070202

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ: British Medical Journal, 327,* 557–560. http://dx.doi.org/10.1136/bmj.327.7414.557

*#Hu, X., Chen, H., & Fu, G. (2012). A repeated lie becomes a truth? The effect of intentional control and training on deception. *Frontiers in Psychology, 3,* 488. http://dx.doi.org/10.3389/fpsyg.2012.00488

*Hu, X., Evans, A., Wu, H., Lee, K., & Fu, G. (2013). An interfering dot-probe task facilitates the detection of mock crime memory in a reaction time (RT)-based concealed information test. *Acta Psychologica, 142,* 278–285. http://dx.doi.org/10.1016/j.actpsy.2012.12.006

Hu, X., Hegeman, D., Landry, E., & Rosenfeld, J. P. (2012). Increasing the number of irrelevant stimuli increases ability to detect countermeasures to the P300-based Complex Trial Protocol for concealed information detection. *Psychophysiology, 49,* 85–95. http://dx.doi.org/10.1111/j.1469-8986.2011.01286.x

Hu, X., Pornpattananangkul, N., & Nusslock, R. (2015). Executive control- and reward-related neural processes associated with the opportunity to engage in voluntary dishonest moral decision making. *Cognitive, Affective & Behavioral Neuroscience, 15,* 475–491. http://dx.doi.org/10.3758/s13415-015-0336-9

#Hu, X., Rosenfeld, J. P., & Bodenhausen, G. V. (2012). Combating automatic autobiographical associations: The effect of instruction and training in strategically concealing information in the autobiographical implicit association test. *Psychological Science, 23,* 1079–1085. http://dx.doi.org/10.1177/0956797612443834

*Hu, X., Wu, H., & Fu, G. (2011). Temporal course of executive control when lying about self- and other-referential information: An ERP study. *Brain Research, 1369,* 149–157. http://dx.doi.org/10.1016/j.brainres.2010.10.106

#Huntjens, R. J., Verschuere, B., & McNally, R. J. (2012). Inter-identity autobiographical amnesia in patients with dissociative identity disorder. *PLoS ONE, 7,* e40580. http://dx.doi.org/10.1371/journal.pone.0040580

Inzlicht, M., & Schmeichel, B. J. (2012). What is ego depletion? Toward a mechanistic revision of the resource model of self-control. *Perspectives on Psychological Science, 7,* 450–463. http://dx.doi.org/10.1177/1745691612454134

Ioannidis, J. P. (2005). Contradicted and initially stronger effects in highly cited clinical research. *Journal of the American Medical Association, 294,* 218–228. http://dx.doi.org/10.1001/jama.294.2.218

Ioannidis, J. P. A. (2006). Journals should publish all "null" results and should sparingly publish "positive" results. *Cancer Epidemiology, Biomarkers & Prevention, 15,* 186. http://dx.doi.org/10.1158/1055-9965.EPI-05-0921

*Ito, A., Abe, N., Fujii, T., Hayashi, A., Ueno, A., Mugikura, S., . . . Mori, E. (2012). The contribution of the dorsolateral prefrontal cortex to the preparation for deception and truth-telling. *Brain Research, 1464,* 43–52. http://dx.doi.org/10.1016/j.brainres.2012.05.004

*Ito, A., Abe, N., Fujii, T., Ueno, A., Koseki, Y., Hashimoto, R., . . . Mori, E. (2011). The role of the dorsolateral prefrontal cortex in deception when remembering neutral and emotional events. *Neuroscience Research, 69,* 121–128. http://dx.doi.org/10.1016/j.neures.2010.11.001

Jensen, A. R. (2006). *Clocking the mind: Mental chronometry and individual differences*. Amsterdam, the Netherlands: Elsevier.

*Jiang, W., Liu, H., Liao, J., Ma, X., Rong, P., Tang, Y., & Wang, W. (2013). A functional MRI study of deception among offenders with antisocial personality disorders. *Neuroscience, 244,* 90–98. http://dx.doi.org/10.1016/j.neuroscience.2013.03.055

*Johnson, R., Jr., Barnhardt, J., & Zhu, J. (2003). The deceptive response: Effects of response conflict and strategic monitoring on the late positive component and episodic memory-related brain activity. *Biological Psychology, 64,* 217–253. http://dx.doi.org/10.1016/j.biopsycho.2003.07.006

Johnson, R., Jr., Barnhardt, J., & Zhu, J. (2004). The contribution of executive processes to deceptive responding. *Neuropsychologia, 42,* 878–901. http://dx.doi.org/10.1016/j.neuropsychologia.2003.12.005

Johnson, R., Jr., Barnhardt, J., & Zhu, J. (2005). Differential effects of practice on the executive processes used for truthful and deceptive responses: An event-related brain potential study. *Cognitive Brain Research, 24,* 386–404. http://dx.doi.org/10.1016/j.cogbrainres.2005.02.011

*Johnson, R., Jr., Henkell, H., Simon, E., & Zhu, J. (2008). The self in conflict: The role of executive processes during truthful and deceptive responses about attitudes. *NeuroImage, 39,* 469–482. http://dx.doi.org/10.1016/j.neuroimage.2007.08.032

Jokinen, A., Santtila, P., Ravaja, N., & Puttonen, S. (2006). Salience of Guilty Knowledge Test items affects accuracy in realistic mock crimes. *International Journal of Psychophysiology, 62,* 175–184. http://dx.doi.org/10.1016/j.ijpsycho.2006.04.004

Jung, C. G. (1910). The association method. *The American Journal of Psychology, 21,* 219–269. http://dx.doi.org/10.2307/1413002

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General, 132,* 47–70. http://dx.doi.org/10.1037/0096-3445.132.1.47

*Kaylor-Hughes, C. J., Lankappa, S. T., Fung, R., Hope-Urwin, A. E., Wilkinson, I. D., & Spence, S. A. (2011). The functional anatomical distinction between truth telling and deception is preserved among people with schizophrenia. *Criminal Behaviour and Mental Health, 21,* 8–20. http://dx.doi.org/10.1002/cbm.785

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching—A review. *Psychological Bulletin, 136,* 849–874. http://dx.doi.org/10.1037/a0019842

Kleinberg, B., & Verschuere, B. (2015). Memory detection 2.0: The first web-based memory detection test. *PLoS ONE, 10,* e0118715. http://dx.doi.org/10.1371/journal.pone.0118715

Kleinberg, B., & Verschuere, B. (in press). The role of motivation to avoid detection in reaction time-based concealed information detection. *Journal of Applied Research in Memory & Cognition*.

Kleinmuntz, B., & Szucko, J. J. (1984). Lie detection in ancient and modern times. A call for contemporary scientific study. *American Psychologist, 39,* 766–776. http://dx.doi.org/10.1037/0003-066X.39.7.766

Klein Selle, N., Verschuere, B., Kindt, M., Meijer, E., & Ben-Shakhar, G. (2016). Orienting versus inhibition in the Concealed Information Test: Different cognitive processes drive different physiological measures. *Psychophysiology, 53,* 579–590. http://dx.doi.org/10.1111/psyp.12583

Kouchaki, M., & Smith, I. H. (2014). The morning morality effect: The influence of time of day on unethical behavior. *Psychological Science, 25,* 95–102. http://dx.doi.org/10.1177/0956797613498099

Krauss, R. M. (1981). Impression formation, impression management, and nonverbal behaviors. Paper presented at Social Cognition: The Ontario Symposium.

*Kubo, K., & Nittono, H. (2009). The role of intention to conceal in the P300-based concealed information test. *Applied Psychophysiology and Biofeedback, 34,* 227–235. http://dx.doi.org/10.1007/s10484-009-9089-y

Lancaster, G. L., Vrij, A., Hope, L., & Waller, B. (2013). Sorting the liars from the truth tellers: The benefits of asking unanticipated questions on lie detection. *Applied Cognitive Psychology, 27,* 107–114. http://dx.doi.org/10.1002/acp.2879

Langleben, D. D., Loughead, J. W., Bilker, W. B., Ruparel, K., Childress, A. R., Busch, S. I., & Gur, R. C. (2005). Telling truth from lie in individual subjects with fast event-related fMRI. *Human Brain Mapping, 26,* 262–272. http://dx.doi.org/10.1002/hbm.20191

Lee, T. M., Lee, T. M., Raine, A., & Chan, C. C. (2010). Lying about the valence of affective pictures: An fMRI study. *PLoS ONE, 5,* e12291. http://dx.doi.org/10.1371/journal.pone.0012291

Legault, L., & Inzlicht, M. (2013). Self-determination, self-regulation, and the brain: Autonomy improves performance by enhancing neuroaffective responsiveness to self-regulation failure. *Journal of Personality and Social Psychology, 105,* 123–138. http://dx.doi.org/10.1037/a0030426

*Leue, A., Lange, S., & Beauducel, A. (2012). "Have you ever seen this face?" - Individual differences and event-related potentials during deception. *Frontiers in Psychology, 3,* 570. http://dx.doi.org/10.3389/fpsyg.2012.00570

Levine, T. R. (2014). Truth-default theory (TDT): A theory of human deception and deception detection. *Journal of Language and Social Psychology, 33,* 378–392. http://dx.doi.org/10.1177/0261927X14535916

Lewin, K. (1951). *Field theory in social science: Selected theoretical papers* (D. Cartwright, Ed.). New York, NY: Harper & Row.

Lieblich, I., Ben-Shakhar, G., & Kugelmass, S. (1976). Validity of the guilty of knowledge technique in a prisoner's sample. *Journal of Applied Psychology, 61,* 89–93. http://dx.doi.org/10.1037/0021-9010.61.1.89

Lieblich, I., Kugelmass, S., & Ben-Shakhar, G. (1970). Efficiency of GSR detection of information as a function of stimulus set size. *Psychophysiology, 6,* 601–608. http://dx.doi.org/10.1111/j.1469-8986.1970.tb02249.x

Lipsey, M. W., & Wilson, D. B. (2001). *Applied social research methods series: Vol. 49. Practical meta analysis*. Thousand Oaks, CA: Sage.

Locker, L., & Pratarelli, M. (1997). Lexical decision and the detection of concealed information. *The Journal of Credibility Assessment and Witness Psychology, 1,* 33–43.

*Lui, M., & Rosenfeld, J. P. (2009). The application of subliminal priming in lie detection: Scenario for identification of members of a terrorist ring. *Psychophysiology, 46,* 889–903. http://dx.doi.org/10.1111/j.1469-8986.2009.00810.x

Luria, A. (1932). *The nature of human conflicts*. New York, NY: Liveright.

Lykken, D. T. (1959). The Gsr in the detection of guilt. *Journal of Applied Psychology, 43,* 385–388. http://dx.doi.org/10.1037/h0046060

*Mameli, F., Tomasini, E., Scelzo, E., Fumagalli, M., Ferrucci, R., Bertolasi, L., & Priori, A. (2013). Lies tell the truth about cognitive dysfunction in essential tremor: An experimental deception study with the guilty knowledge task. *Journal of Neurology, Neurosurgery, & Psychiatry, 84,* 1008–1013. http://dx.doi.org/10.1136/jnnp-2012-304674

*Marchewka, A., Jednorog, K., Falkiewicz, M., Szeszkowski, W., Grabowska, A., & Szatkowska, I. (2012). Sex, lies and fMRI—Gender differences in neural basis of deception. *PLoS ONE, 7,* e43076. http://dx.doi.org/10.1371/journal.pone.0043076

Marston, W. M. (1920). Reaction-time symptoms of deception. *Journal of Experimental Psychology, 3,* 72–87. http://dx.doi.org/10.1037/h0067963

*Matsuda, I., Nittono, H., & Allen, J. J. B. (2013). Detection of concealed information by P3 and frontal EEG asymmetry. *Neuroscience Letters, 537,* 55–59. http://dx.doi.org/10.1016/j.neulet.2013.01.029

*Matsuda, I., Nittono, H., Hirota, A., Ogawa, T., & Takasawa, N. (2009). Event-related brain potentials during the standard autonomic-based concealed information test. *International Journal of Psychophysiology, 74,* 58–68. http://dx.doi.org/10.1016/j.ijpsycho.2009.07.004

McCornack, S., Morrison, K., Paik, J. E., Wisner, A. M., & Zhu, X. (2014). Information manipulation theory 2: A propositional theory of deceptive discourse production. *Journal of Language and Social Psychology, 33,* 348–377. http://dx.doi.org/10.1177/0261927X14534656

Meijer, E. H., Selle, N. K., Elber, L., & Ben-Shakhar, G. (2014). Memory detection with the Concealed Information Test: A meta analysis of skin conductance, respiration, heart rate, and P300 data. *Psychophysiology, 51,* 879–904. http://dx.doi.org/10.1111/psyp.12239

*Meijer, E. H., Smulders, F. T. Y., Merckelbach, H. L. G. J., & Wolf, A. G. (2007). The P300 is sensitive to concealed face recognition. *International Journal of Psychophysiology, 66,* 231–237. http://dx.doi.org/10.1016/j.ijpsycho.2007.08.001

Meijer, E. H., Verschuere, B., Gamer, M., Merckelbach, H., & Ben-Shakhar, G. (2016). Deception detection with behavioral, autonomic, and neural measures: Conceptual and methodological considerations that warrant modesty. *Psychophysiology, 53,* 593–604. http://dx.doi.org/10.1111/psyp.12609

*#Mertens, R., & Allen, J. J. B. (2008). The role of psychophysiology in forensic assessments: Deception detection, ERPs, and virtual reality

mock crime scenarios. *Psychophysiology, 45,* 286–298. http://dx.doi.org/10.1111/j.1469-8986.2007.00615.x

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41,* 49–100. http://dx.doi.org/10.1006/cogp.1999.0734

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences, 7,* 134–140. http://dx.doi.org/10.1016/S1364-6613(03)00028-7

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods, 7,* 105–125. http://dx.doi.org/10.1037/1082-989X.7.1.105

Muraven, M., & Slessareva, E. (2003). Mechanisms of self-control failure: Motivation and limited resources. *Personality and Social Psychology Bulletin, 29,* 894–906. http://dx.doi.org/10.1177/0146167203029007008

National Research Council. (2003). *The polygraph and lie detection. Committee to review the scientific evidence on the Polygraph. Division of Behavioral and Social Sciences and Education*. Washington, DC: The National Academic Press.

*Noordraven, E., & Verschuere, B. (2013). Predicting the Sensitivity of the Reaction Time-based Concealed Information Test. *Applied Cognitive Psychology, 27,* 328–335. http://dx.doi.org/10.1002/acp.2910

*Nose, I., Murai, J., & Taira, M. (2009). Disclosing concealed information on the basis of cortical activations. *NeuroImage, 44,* 1380–1386. http://dx.doi.org/10.1016/j.neuroimage.2008.11.002

Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific utopia II. Restructuring incentives and practices to promote truth over publishability. *Perspectives on Psychological Science, 7,* 615–631. http://dx.doi.org/10.1177/1745691612459058

Panasiti, M. S., Pavone, E. F., Mancini, A., Merla, A., Grisoni, L., & Aglioti, S. M. (2014). The motor cost of telling lies: Electrocortical signatures and personality foundations of spontaneous deception. *Social Neuroscience, 9,* 573–589. http://dx.doi.org/10.1080/17470919.2014.934394

Pashler, H., & Wagenmakers, E. J. (2012). Editors' introduction to the special section on replicability in psychological science: A crisis of confidence? *Perspectives on Psychological Science, 7,* 528–530. http://dx.doi.org/10.1177/1745691612465253

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience, 9,* 148–158. http://dx.doi.org/10.1038/nrn2317

*Pfister, R., Foerster, A., & Kunde, W. (2014). Pants on fire: The electrophysiological signature of telling a lie. *Social Neuroscience, 9,* 562–572.

*Priori, A., Mameli, F., Cogiamanian, F., Marceglia, S., Tiriticco, M., Mrakic-Sposta, S., . . . Sartori, G. (2008). Lie-specific involvement of dorsolateral prefrontal cortex in deception. *Cerebral Cortex, 18,* 451–455. http://dx.doi.org/10.1093/cercor/bhm088

Proverbio, A. M., Vanutelli, M. E., & Adorni, R. (2013). Can you catch a liar? How negative emotions affect brain responses when lying or telling the truth. *PLoS ONE, 8,* e59383. http://dx.doi.org/10.1371/journal.pone.0059383

Reid, J. E. (1947). A revised questioning technique in lie-detection tests. *Journal of Criminal Law and Criminology (1931–1951), 37,* 542–547. http://dx.doi.org/10.2307/1138979

Ridderinkhof, K. R., van den Wildenberg, W. P., Wijnen, J., & Burle, B. (2004). Response inhibition in conflict tasks is revealed in delta plots. In M. I. Posner (Ed.), *Cognitive Neuroscience of Attention* (pp. 369–377). New York, NY: Guilford Press.

*Rosenfeld, J. P., Hu, X., & Pederson, K. (2012). Deception awareness improves P300-based deception detection in concealed information tests. *International Journal of Psychophysiology, 86,* 114–121. http://dx.doi.org/10.1016/j.ijpsycho.2012.06.007

Rosenfeld, J. P., Labkovsky, E., Winograd, M., Lui, M. A., Vandenboom, C., & Chedid, E. (2008). The Complex Trial Protocol (CTP): A new, countermeasure-resistant, accurate, P300-based method for detection of concealed information. *Psychophysiology, 45,* 906–919. http://dx.doi.org/10.1111/j.1469-8986.2008.00708.x

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86,* 638–641. http://dx.doi.org/10.1037/0033-2909.86.3.638

Rouder, J. N., & Morey, R. D. (2005). Relational and arelational confidence intervals: A comment on Fidler, Thomason, Cumming, Finch, and Leeman (2004). *Psychological Science, 16,* 77–79. http://dx.doi.org/10.1111/j.0956-7976.2005.00783.x

*Sai, L., Lin, X., Hu, X., & Fu, G. (2014). Detecting concealed information using feedback related event-related brain potentials. *Brain and Cognition, 90,* 142–150. http://dx.doi.org/10.1016/j.bandc.2014.06.012

*Sartori, G., Agosta, S., Zogmaister, C., Ferrara, S. D., & Castiello, U. (2008). How to accurately detect autobiographical events. *Psychological Science, 19,* 772–780. http://dx.doi.org/10.1111/j.1467-9280.2008.02156.x

Schooler, J. (2011). Unpublished results hide the decline effect. *Nature, 470,* 437. http://dx.doi.org/10.1038/470437a

Seymour, T. L., & Kerlin, J. R. (2008). Successful detection of verbal and visual concealed knowledge using an RT-based paradigm. *Applied Cognitive Psychology, 22,* 475–490. http://dx.doi.org/10.1002/acp.1375

Seymour, T. L., Seifert, C. M., Shafto, M. G., & Mosmann, A. L. (2000). Using response time measures to assess "guilty knowledge." *Journal of Applied Psychology, 85,* 30–37. http://dx.doi.org/10.1037/0021-9010.85.1.30

Shalvi, S., Eldar, O., & Bereby-Meyer, Y. (2012). Honesty requires time (and lack of justifications). *Psychological Science, 23,* 1264–1270. http://dx.doi.org/10.1177/0956797612443835

*Sheridan, R. S., & Flowers, K. A. (2010). Reaction times and deception: The lying constant. *International Journal of Psychological Studies, 2,* 41–51. http://dx.doi.org/10.5539/ijps.v2n2p41

Sip, K. E., Carmel, D., Marchant, J. L., Li, J., Petrovic, P., Roepstorff, A., . . . Frith, C. D. (2013). When Pinocchio's nose does not grow: Belief regarding lie-detectability modulates production of deception. *Frontiers in Human Neuroscience, 7,* 16. http://dx.doi.org/10.3389/fnhum.2013.00016

Sip, K. E., Roepstorff, A., McGregor, W., & Frith, C. D. (2008). Detecting deception: The scope and limits. *Trends in Cognitive Sciences, 12,* 48–53. http://dx.doi.org/10.1016/j.tics.2007.11.008

Smith, G. T., Fischer, S., & Fister, S. M. (2003). Incremental validity principles in test construction. *Psychological Assessment, 15,* 467–477. http://dx.doi.org/10.1037/1040-3590.15.4.467

*Spence, S. A., Farrow, T. F. D., Herford, A. E., Wilkinson, I. D., Zheng, Y., & Woodruff, P. W. R. (2001). Behavioural and functional anatomical correlates of deception in humans. *NeuroReport, 12,* 2849–2853. http://dx.doi.org/10.1097/00001756-200109170-00019

Spence, S. A., Hunter, M. D., Farrow, T. F., Green, R. D., Leung, D. H., Hughes, C. J., & Ganesan, V. (2004). A cognitive neurobiological account of deception: Evidence from functional neuroimaging. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences, 359,* 1755–1762. http://dx.doi.org/10.1098/rstb.2004.1555

Spence, S. A., Kaylor-Hughes, C. J., Brook, M. L., Lankappa, S. T., & Wilkinson, I. D. (2008). 'Munchausen's syndrome by proxy' or a 'miscarriage of justice'? An initial application of functional neuroimaging to the question of guilt versus innocence. *European Psychiatry, 23,* 309–314. http://dx.doi.org/10.1016/j.eurpsy.2007.09.001

Sporer, S. L. (2016). Deception and cognitive load: Expanding our horizon with a working memory model. *Frontiers in Psychology, 7*. http://dx.doi.org/10.3389/fpsyg.2016.00420

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

Sterling, T. D. (1959). Publication decisions and their possible effects on inferences drawn from tests of significance—Or vice versa. *Journal of the American Statistical Association, 54,* 30–34.

Sterling, T. D., Rosenbaum, W., & Weinkam, J. (1995). Publication decisions revisited: The effect of the outcome of statistical tests on the decision to publish and vice versa. *The American Statistician, 49,* 108–112.

Sterne, J. A., Gavaghan, D., & Egger, M. (2000). Publication and related bias in meta-analysis: Power of statistical tests and prevalence in the literature. *Journal of Clinical Epidemiology, 53,* 1119–1129. http://dx.doi.org/10.1016/S0895-4356(00)00242-0

*Suchotzki, K., Crombez, G., Debey, E., van Oorsouw, K., & Verschuere, B. (2015). In vino veritas? Alcohol, response inhibition and lying. *Alcohol and Alcoholism, 50,* 74–81. http://dx.doi.org/10.1093/alcalc/agu079

*Suchotzki, K., Crombez, G., Smulders, F. T., Meijer, E., & Verschuere, B. (2015). The cognitive mechanisms underlying deception: An event-related potential study. *International Journal of Psychophysiology, 95,* 395–405. http://dx.doi.org/10.1016/j.ijpsycho.2015.01.010

*Suchotzki, K., Verschuere, B., Crombez, G., & De Houwer, J. (2013). Reaction time measures in deception research: Comparing the effects of irrelevant and relevant stimulus-response compatibility. *Acta Psychologica, 144,* 224–231. http://dx.doi.org/10.1016/j.actpsy.2013.06.014

*Suchotzki, K., Verschuere, B., Peth, J., Crombez, G., & Gamer, M. (2015). Manipulating item proportion and deception reveals crucial dissociation between behavioral, autonomic, and neural indices of concealed information. *Human Brain Mapping, 36,* 427–439. http://dx.doi.org/10.1002/hbm.22637

Tabatabaeian, M., Dale, R., & Duran, N. D. (2015). Self-serving dishonest decisions can show facilitated cognitive dynamics. *Cognitive Processing, 16,* 291–300. http://dx.doi.org/10.1007/s10339-015-0660-6

Thompson, R. F. (2009). Habituation: A history. *Neurobiology of Learning and Memory, 92,* 127–134. http://dx.doi.org/10.1016/j.nlm.2008.07.011

Thompson, S. G., & Higgins, J. P. (2002). How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine, 21,* 1559–1573. http://dx.doi.org/10.1002/sim.1187

Tomash, J. J., & Reed, P. (2015). Using conditioning to elicit skin conductance responses to deception. *Learning and Motivation, 49,* 31–37. http://dx.doi.org/10.1016/j.lmot.2015.02.002

Turner, R. M., Bird, S. M., & Higgins, J. P. (2013). The impact of study size on meta-analyses: Examination of underpowered studies in Cochrane reviews. *PLoS ONE, 8,* e59202. http://dx.doi.org/10.1371/journal.pone.0059202

#Uncapher, M. R., Boyd-Meredith, J. T., Chow, T. E., Rissman, J., & Wagner, A. D. (2015). Goal-directed modulation of neural memory patterns: Implications for fMRI-based memory detection. *The Journal of Neuroscience, 35,* 8531–8545. http://dx.doi.org/10.1523/JNEUROSCI.5145-14.2015

*Van Bockstaele, B., Verschuere, B., Moens, T., Suchotzki, K., Debey, E., & Spruyt, A. (2012). Learning to lie: Effects of practice on the cognitive cost of lying. *Frontiers in Psychology, 3,* 526. http://dx.doi.org/10.3389/fpsyg.2012.00526

Van Bockstaele, B., Wilhelm, C., Meijer, E., Debey, E., & Verschuere, B. (2015). When deception becomes easy: The effects of task switching and goal neglect on the truth proportion effect. *Frontiers in Psychology, 6,* 1666. http://dx.doi.org/10.3389/fpsyg.2015.01666

Vandierendonck, A., Liefooghe, B., & Verbruggen, F. (2010). Task switching: Interplay of reconfiguration and interference control. *Psychological Bulletin, 136,* 601–626. http://dx.doi.org/10.1037/a0019791

van Hooff, J. C., Sargeant, E., Foster, J. K., & Schmand, B. A. (2009). Identifying deliberate attempts to fake memory impairment through the combined use of reaction time and event-related potential measures. *International Journal of Psychophysiology, 73,* 246–256. http://dx.doi.org/10.1016/j.ijpsycho.2009.04.002

*Varga, M., Visu-Petra, G., Miclea, M., & Visu-Petra, L. (2015). The "good cop, bad cop" effect in the RT-based concealed information test: Exploring the effect of emotional expressions displayed by a virtual investigator. *PLoS ONE, 10,* e0116087. http://dx.doi.org/10.1371/journal.pone.0116087

*Vartanian, O., Kwantes, P., & Mandel, D. R. (2012). Lying in the scanner: Localized inhibition predicts lying skill. *Neuroscience Letters, 529,* 18–22. http://dx.doi.org/10.1016/j.neulet.2012.09.019

*Vartanian, O., Kwantes, P. J., Mandel, D. R., Bouak, F., Nakashima, A., Smith, I., & Lam, Q. (2013). Right inferior frontal gyrus activation as a neural marker of successful lying. *Frontiers in Human Neuroscience, 7,* 616. http://dx.doi.org/10.3389/fnhum.2013.00616

*Vendemia, J. M. C., Buzan, R. F., & Green, E. P. (2005). Practice effects, workload, and reaction time in deception. *The American Journal of Psychology, 118,* 413–429.

*Vendemia, J. M. C., Buzan, R. F., & Simon-Dack, S. L. (2005). Reaction time of motor responses in two-stimulus paradigms involving deception and congruity with varying levels of difficulty. *Behavioural Neurology, 16,* 25–36. http://dx.doi.org/10.1155/2005/804026

*Verschuere, B., Crombez, G., Degrootte, T., & Rosseel, Y. (2010). Detecting concealed information with reaction times: Validity and comparison with the polygraph. *Applied Cognitive Psychology, 24,* 991–1002. http://dx.doi.org/10.1002/acp.1601

Verschuere, B., Crombez, G., Koster, E. H., Van Bockstaele, B., & De Clercq, A. (2007). Startling secrets: Startle eye blink modulation by concealed crime information. *Biological Psychology, 76,* 52–60. http://dx.doi.org/10.1016/j.biopsycho.2007.06.001

Verschuere, B., & De Houwer, J. (2011). Detecting concealed information in less than a second: Response latency-based measures. In B. Verschuere, G. Ben-Shakhar, & E. Meijer (Eds.), *Memory detection: Theory and application of the Concealed Information Test* (pp. 46–62). Cambridge, UK: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511975196.004

*Verschuere, B., Kleinberg, B., & Theocharidou, K. (2015). RT-based memory detection: Item saliency effects in the single-probe and the multiple-probe protocol. *Journal of Applied Research in Memory & Cognition, 4,* 59–65. http://dx.doi.org/10.1016/j.jarmac.2015.01.001

Verschuere, B., & Meijer, E. (2014). What's on your mind? Detecting concealed information. *European Psychologist, 19,* 162–171. http://dx.doi.org/10.1027/1016-9040/a000194

*#Verschuere, B., Prati, V., & Houwer, J. D. (2009). Cheating the lie detector: Faking in the autobiographical Implicit Association Test. *Psychological Science, 20,* 410–413. http://dx.doi.org/10.1111/j.1467-9280.2009.02308.x

*Verschuere, B., Rosenfeld, J. P., Winograd, M. R., Labkovsky, E., & Wiersema, R. (2009). The role of deception in P300 memory detection. *Legal and Criminological Psychology, 14,* 253–262. http://dx.doi.org/10.1348/135532508X384184

*Verschuere, B., Schuhmann, T., & Sack, A. T. (2012). Does the inferior frontal sulcus play a functional role in deception? A neuronavigated theta-burst transcranial magnetic stimulation study. *Frontiers in Human Neuroscience, 6,* 284. http://dx.doi.org/10.3389/fnhum.2012.00284

Verschuere, B., & Shalvi, S. (2014). The truth comes naturally! Does it? *Journal of Language and Social Psychology, 33,* 417–423. http://dx.doi.org/10.1177/0261927X14535394

*Verschuere, B., Spruyt, A., Meijer, E. H., & Otgaar, H. (2011). The ease of lying. *Consciousness and Cognition, 20,* 908–911. http://dx.doi.org/10.1016/j.concog.2010.10.023

Verschuere, B., Suchotzki, K., & Debey, E. (2015). Detecting deception through reaction times. In P. A. Granhag, A. Vrij, & B. Verschuere (Eds.), *Deception detection: Current challenges and new approaches*. Chichester, UK: Wiley, Ltd.

Vidal, F., Hasbroucq, T., Grapperon, J., & Bonnet, M. (2000). Is the 'error negativity' specific to errors? *Biological Psychology, 51,* 109–128. http://dx.doi.org/10.1016/S0301-0511(99)00032-0

Vincent, A., & Furedy, J. J. (1992). Electrodermal differentiation of deception: Potentially confounding and influencing factors. *International Journal of Psychophysiology, 13,* 129–136. http://dx.doi.org/10.1016/0167-8760(92)90052-D

*Visu-Petra, G., Miclea, M., Bus, I., & Visu-Petra, L. (2014). Detecting concealed information: The role of individual differences in executive functions and social desirability. *Psychology, Crime & Law, 20,* 20–36. http://dx.doi.org/10.1080/1068316X.2012.736509

*Visu-Petra, G., Miclea, M., & Visu-Petra, L. (2012). Reaction time-based detection of concealed information in relation to individual differences in executive functioning. *Applied Cognitive Psychology, 26,* 342–351. http://dx.doi.org/10.1002/acp.1827

*Visu-Petra, G., Varga, M., Miclea, M., & Visu-Petra, L. (2013). When interference helps: Increasing executive load to facilitate deception detection in the concealed information test. *Frontiers in Psychology, 4,* 146. http://dx.doi.org/10.3389/fpsyg.2013.00146

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

Vrij, A., Fisher, R. P., & Blank, H. (2015). A cognitive approach to lie detection: A meta-analysis. *Legal and Criminological Psychology, 22,* 1–21. http://dx.doi.org/10.1111/lcrp.12088

Vrij, A., Fisher, R., Mann, S., & Leal, S. (2006). Detecting deception by manipulating cognitive load. *Trends in Cognitive Sciences, 10,* 141–142. http://dx.doi.org/10.1016/j.tics.2006.02.003

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

Vrij, A., Leal, S., Granhag, P. A., Mann, S., Fisher, R. P., Hillman, J., & Sperry, K. (2009). Outsmarting the liars: The benefit of asking unanticipated questions. *Law and Human Behavior, 33,* 159–166. http://dx.doi.org/10.1007/s10979-008-9143-y

Vrij, A., Mann, S. A., Fisher, R. P., Leal, S., Milne, R., & Bull, R. (2008). Increasing cognitive load to facilitate lie detection: The benefit of recalling an event in reverse order. *Law and Human Behavior, 32,* 253–265. http://dx.doi.org/10.1007/s10979-007-9103-y

Vrij, A., Semin, G. R., & Bull, R. (1996). Insight into behavior displayed during deception. *Human Communication Research, 22,* 544–562. http://dx.doi.org/10.1111/j.1468-2958.1996.tb00378.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. http://dx.doi.org/10.1016/j.newideapsych.2014.03.001

Walczyk, J. J., Roper, K. S., Seemann, E., & Humphrey, A. M. (2003). Cognitive mechanisms underlying lying to questions: Response time as a cue to deception. *Applied Cognitive Psychology, 17,* 755–774. http://dx.doi.org/10.1002/acp.914

Warmelink, L., Vrij, A., Mann, S., Jundi, S., & Granhag, P. A. (2012). The effect of question expectedness and experience on lying about intentions. *Acta Psychologica, 141,* 178–183. http://dx.doi.org/10.1016/j.actpsy.2012.07.011

*Williams, E. J. (2012). *Lies and cognition: How do we tell lies and can we detect them?* (Unpublished doctoral dissertation). Cardiff University, Cardiff, Wales.

*Williams, E. J., Bott, L. A., Patrick, J., & Lewis, M. B. (2013). Telling lies: The irrepressible truth? *PLoS ONE, 8*(4), e60713. http://dx.doi.org/10.1371/journal.pone.0060713

*Wu, H., Hu, X., & Fu, G. (2009). Does willingness affect the N2–P3 effect of deceptive and honest responses? *Neuroscience Letters, 467,* 63–66. http://dx.doi.org/10.1016/j.neulet.2009.10.002

Zhang, Z., Xu, X., & Ni, H. (2013). Small studies may overestimate the effect sizes in critical care meta-analyses: A meta-epidemiological study. *Critical Care (London, England), 17*(1), R2. http://dx.doi.org/10.1186/cc11919

*Zhao, M., Zheng, C., & Zhao, C. (2012). A new approach for concealed information identification based on ERP assessment. *Journal of Medical Systems, 36,* 2401–2409. http://dx.doi.org/10.1007/s10916-011-9707-0

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. *Advances in Experimental Social Psychology, 14,* 59.

Received January 8, 2015
Revision received October 24, 2016
Accepted November 3, 2016 ■