Cognition 114 (2010) 105–110

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Brief article

A truth that's told with bad intent: An ERP study of deception

Ricardo E. Carrión [a,*], Julian P. Keenan [b], Natalie Sebanz [c]

[a] Department of Psychiatry, Zucker Hillside Hospital, Long Island, NY, United States
[b] Montclair State University, NJ, United States
[c] Donders Institute for Brain, Cognition, and Behaviour, Radboud University Nijmegen, The Netherlands

ARTICLE INFO

Article history:
Received 15 October 2008
Revised 18 May 2009
Accepted 22 May 2009

Keywords:
Deception
Event-related potentials
Cognitive control
Mentalizing
Social interaction

ABSTRACT

Human social cognition critically relies on the ability to deceive others. However, the cognitive and neural underpinnings of deception are still poorly understood. Why does lying place increased demands on cognitive control? The present study investigated whether cognitive control processes during deception are recruited due to the need to inhibit a tendency to state the truth, or reflect deceptive intent more generally. We engaged participants in a face-to-face interaction game and examined event-related brain potentials (ERPs) while participants lied and told the truth with or without deceptive intention. The same medial frontal negative deflection (N450) occurred when participants lied and when they told the truth with deceptive intent. This suggests that the main challenge of lying is not to inhibit a tendency to state the truth. Rather, the challenge is to handle the cognitive conflict resulting from the need to keep others' mental states in mind while deceiving them.

1. Introduction

Humans deceive each other in many different ways: intentionally withholding information, providing false information, and, sometimes, telling the truth in an attempt to mislead others, which is effective when we have reason to assume that others are not going to believe us (Happé, 1994). Research on the cognitive and neural processes underlying deception has focused on how individuals make counterfactual statements, commonly known as lying (for a review, see Spence et al., 2004). Neuroimaging and ERP studies have shown that lying is associated with increased demands on cognitive control, which is generally thought to be critical for allocating mental resources to processing task-relevant information, inhibiting predominant response tendencies, and guiding behavior in situations involving response conflict (e.g., Botvinick, Braver, Barch, Carter, & Cohen, 2001). Lie responses, in contrast to truthful responses, have been associated with increased activation in several prefrontal regions linked to cognitive control, including dorsolateral prefrontal cortex and anterior cingulate cortex (Asbe, Suzuki, Mori, Itoh, & Fujii, 2007; Ganis, Kosslyn, Stose, Thompson, & Yurgelun-Todd, 2003; Nuñez, Casey, Egner, Hare, & Hirsch, 2005; Spence et al., 2001).

To the best of our knowledge, it is unknown whether other forms of deception, such as telling the truth with deceptive intent (also known as a double bluff, cf. Happé (1994)), place similar demands on cognitive control processes. On the one hand, previous research suggests that increased demands on cognitive control should selectively stem from making counterfactual statements, because only when making counterfactual statements does one need to inhibit a tendency to state the truth (Asbe et al., 2007; Ganis et al., 2003; Johnson, Barnhardt, & Zhu, 2003, 2005; Kozel, Padgett, & George, 2004; Nuñez et al., 2005; Spence et al., 2001). On the other hand, it is conceivable that lying is challenging in terms of cognitive control because an individual must hold two mental states (i.e., their own

* Corresponding author. Address: Zucker Hillside Hospital, Psychiatry Research, 75-59 263rd Street, Glen Oaks, NY 11004, United States. Tel.: +1 718 470 8878; fax: +1 718 470 8131.
E-mail address: RCarrion@NSHS.edu (Ricardo E. Carrión).

0010-0277/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2009.05.014

and that of another) in mind (Blakemore, Winston, & Fruth, 2004; Proust, 2007). Our tendency to take an intentional stance (Dennett, 1987) and readily attribute mental states to others (Frith & Frith, 2003) may create difficulties in keeping our own and others' perspective apart, thus increasing the need for cognitive control.

Whereas the first account suggests that cognitive control processes are selectively recruited when people state the opposite of what is true, the latter account predicts that the deceiver's intention to deceive should be critical, regardless of whether factual or counterfactual statements are being made. For example, misleading others in a card game by claiming that we are holding an excellent set of cards when this is indeed the case may be as challenging in terms of cognitive control as pretending that our cards are bad. This has not been systematically investigated, because previous studies on deception have focused on the comparison between factual and counterfactual statements. Thus, the main aim of the present study was to focus on the intentional context in which truthful or counterfactual statements are being made, in order to investigate whether the intention to deceive requires the recruitment of cognitive control processes.

A second aim was to create an experimental paradigm that would allow us to study spontaneous deception in a face-to-face interaction. As pointed out in several recent reviews (Blakemore et al., 2004; Sip, Roepstorff, McGregor, & Frith, 2008; Spence et al., 2004), previous cognitive neuroscience studies of deception have been limited in that a single person was instructed to make counterfactual statements, e.g., answering with opposites to autobiographical questions. By embedding deception in a real-time social interaction, we sought to determine whether spontaneous deception follows the same mechanisms as the more engineered deception studied before (cf. Blakemore et al., 2004).

2. The study

Even-related potentials (ERPs) were measured while participants engaged in an interactive face-to-face deception game that involved passing on factual or false/counterfactual information. The ERP analyses focused on a medial frontal negative deflection also known as the N450, a component sensitive to cognitive control demands and conflict processing that is thought to reflect activity in anterior cingulate cortex (West, Jakubek, Wymbs, Perry, & Moore, 2005; West, Bowry, & McConville, 2004). Participants saw a circle or square on the forehead of a person sitting opposite them and tried to deceive the other about the identity of the figure (see Fig. 1). In an ''instructed condition'', a cue indicated whether participants should lie or tell the truth. This allowed us to compare factual vs. counterfactual statements, as in previous studies. In a second, critical ''choice condition'', participants could choose whether to lie or tell the truth. In this condition, participants were led to strategically tell the truth in order to deceive the other. Thus, this condition allowed us to investigate whether telling the truth with ''bad intent'' recruits cognitive control processes to a similar extent as lying.

Our main prediction was that the medial frontal negativity (MFN) would be sensitive to deceptive intent. Specifically, we expected this component to be more pronounced when people told the truth in an attempt to deceive others (choice condition) compared to when people stated the truth without deceptive intent (instructed condition). To further investigate whether the need to keep one's own and the other's perspective apart recruits cognitive control during deception, we also obtained a measure of mentalizing skills in participants. If keeping one's own and another's mental representation apart plays a role in cognitive control during deception, the MFN amplitude should be correlated with mentalizing ability. Furthermore, mentalizing ability may affect individuals' success at deceiving their opponent.

3. Methods

3.1. Participants

Eleven right-handed undergraduates with normal vision (age range 18–22 yrs) gave written informed consent

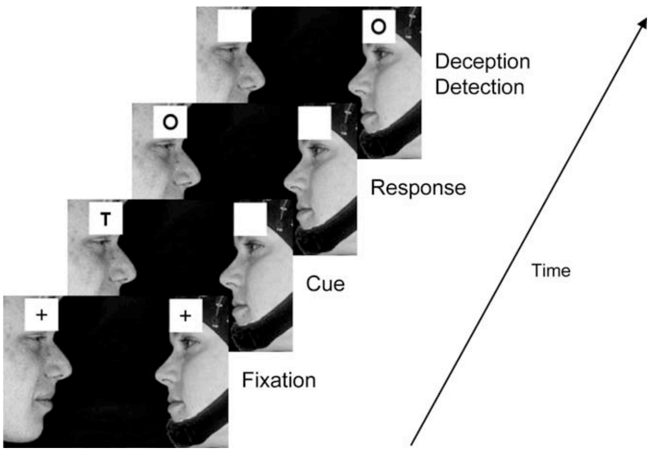

Fig. 1. Example of a 'Truth Instructed' trial, and a 'Truth Chosen' trial.

to participate in the experiment, which was approved by the Rutgers University Institutional Review Board.

3.2. Stimuli and procedure

Participants were fitted with an EEG cap and seated directly facing a confederate at a distance of 1 m. With the help of elastic headbands, a black cardboard square was affixed to the center of each person's head to allow the experimental stimuli to be projected clearly on the forehead (just above the eyes). The stimuli were projected by two video projectors placed behind the participant and confederate.

Participants were asked to play a deception game with the confederate. They were assigned the role of deceiver, while the confederate was assigned the role of opponent. The deceiver saw a circle or square appear on the opponent's forehead that was only visible to them. The deceiver's task was to successfully deceive the opponent about the identity of the stimulus as often as possible during the game. The opponent's task was to correctly guess the identity of the stimulus on their forehead as often as possible. Participants were given a basic monetary fee ($10) for participating and the opportunity to earn additional money ($30) based on their deception success.

In the instructed condition, a cue ("L" or T") instructed the participant to lie or tell the truth. Deceivers were told that the computer randomly selected the cues, and that this would give them a fair chance to beat their opponent. Thus, on instructed truth trials, participants passed on the correct information without intending to deceive the opponent. Instead, they expected that the randomness of the truth and lie instructions would lead to success. In contrast, in the choice condition, a cue ("?") prompted the participant to decide whether to lie or tell the truth. In this condition, participants telling the truth were highly likely to do so with the intention to deceive the other, engaging in *Deceptive Truth*. Importantly, deceivers on both the instructed truth and chosen truth trials saw the same stimuli and performed the same actions. The only difference was whether or not they had the intention to deceive the other while telling the truth.

At the start of each trial a fixation cross was presented (100 ms), which was seen by both the deceiver and the confederate. This was followed by a cue (?, T, or L) for 1000 ms, which was only seen by the deceiver. Following the cue, the deceiver saw either a white circle or square (377×377 pixels, up to 600 ms). The deceiver was instructed to implement their intention as soon as they saw the circle or square by pressing the corresponding key for 'circle' and 'square' on their keyboard. Following the button press, the corresponding figure appeared on the deceiver's forehead. The opponent was instructed to accept or reject the deceiver's choice by pressing the corresponding key for 'circle' and 'square' on their keyboard. This decision was based on what the opponent believed was the stimulus the deceiver had originally seen on their forehead. If the opponent correctly identified the stimulus that the deceiver had seen, 'Match!' appeared on the foreheads' of the deceiver and confederate. Otherwise, 'Liar Wins!' was displayed. Participants completed two blocks of each condition, with each block containing 40 randomized trials. The two blocks of the instructed condition were performed first by all participants. This was done so that participants would not tell the truth with deceptive intent after having adopted this strategy in the choice condition.

Following the experiment, participants' ability to infer others' mental states was measured using Baron-Cohen's test of reading the mind in the eyes (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001), which involves pairing mental states with emotional expressions.

3.3. Electrophysiological recordings

EEG activity was recorded through a DC-based system (Synamps, Compumedics Neuroscan, Inc., El Paso, TX, USA) from a 64-channel Ag/AgCl cap. Electrode positioning in the cap followed the International 10–20 system. The vertical electrooculogram (EOG) was recorded with electrodes placed on the infraorbital and supraorbital foramina of the left eye, and all electrodes were referenced to the left and right mastoids. Maximum impedance at any electrode was 5 KΩ. EEG and EOG were recorded continuously with a sampling rate of 1000 Hz (band-pass .15–200 Hz). Epochs were made of each participant's continuous EEG recording at 100 ms pre-stimulus to 600 ms after stimulus onset (i.e., circle or square on the confederate's forehead). Ocular artifacts were reduced by using Gratton's Eye-movement correction algorithm (Gratton, Coles, & Donchin, 1983). Trials with excessive eye or body movement (voltage exceeding 100 μV at any scalp electrode) were excluded. The ERPs were digitally band-pass filtered between 1 and 10 Hz for presentation purposes. Grand-average (mean waveforms of all participants) ERPs to each trial type were created by averaging the subject-averaged ERPs. The following sites were chosen for statistical analysis: C1, C2, C3, C4, C5, C6, CZ, F1, F2, F3, F4, F5, F6, F7, F8, FZ, P1, P2, P3, P4, P5, P6, P7, P8, PZ, T8, T7. Statistical effects of the corresponding ERP data were tested in an analysis of variance including the repeated measure factors Trial Type (Truth Instructed, Lie Instructed, Truth Chosen, Lie Chosen), Caudality (anterior, central, posterior), and Laterality (left, middle, right). Interactions involving caudality/laterality factors were controlled using vector-normalized amplitudes (McCarthy & Wood, 1985). MFN amplitude was defined as the mean amplitude within a post-stimulus time window of 400–500 ms. Where appropriate, the corrected p values are presented along with the uncorrected degrees of freedom and the Greenhouse–Geisser epsilon (ε) (Greenhouse & Geisser, 1959; Picton et al., 2000). Post-hoc comparisons were computed with Fisher's Protected Least Significant Difference.

4. Results

RTs (Table 1): RTs were slower on lie trials compared to truth trials in the instructed condition ($t\ (10) = 3.08$, $p < .05$), replicating previous findings (Spence et al., 2004). This was not the case in the choice condition ($p = .42$), likely because participants in the choice condition engaged in deception on truth and lie trials. As shown in

Table 1
Mean RTs and standard deviations in the four conditions.

	Instructed *M* (SD) (ms)	Chosen *M* (SD) (ms)
Lie	834 (201)	717 (200)
Truth	769 (190)	700 (203)

Fig. 2a, the ERP analyses revealed a negative deflection about 450 ms post-stimulus that was significantly more pronounced on all deceptive trial types (instructed lie, chosen lie, chosen truth) compared to the instructed truth trials. A repeated measures ANOVA on the MFN amplitude with the factor trial type (instructed truth, instructed lie, chosen truth, chosen lie) revealed a significant effect of trial type, $F(3,96) = 3.79$, $p < .01$, $\varepsilon = .73$, confirming that deceptive intent was associated with increased conflict. The topography of the component was maximal over frontal sites (trial type X caudality: $F(6,192) = 3.78$, $p < .01$, $\varepsilon = .64$) and did not differ reliably for instructed lie, chosen lie, or chosen truth trials (Fig. 2b). Planned comparisons revealed significant differences between the instructed truth trials and all other trial types (all $p < .05$). Instructed lie trials and chosen lie trials were not significantly different from each other ($p = .68$). Chosen truth trials were not significantly different from either chosen lie trials ($p = .94$) or instructed lie trials ($p = .72$). These results demonstrate that telling the truth with a deceptive intention created a similar demand on cognitive control processes as lying.

Correlation analyses between participants' mind in the eyes (MIE) score ($M = 23.5$, $SD = 4.8$) and the MFN amplitude (electrode site: FCz) revealed that the better participants were in the MIE test, the more pronounced was

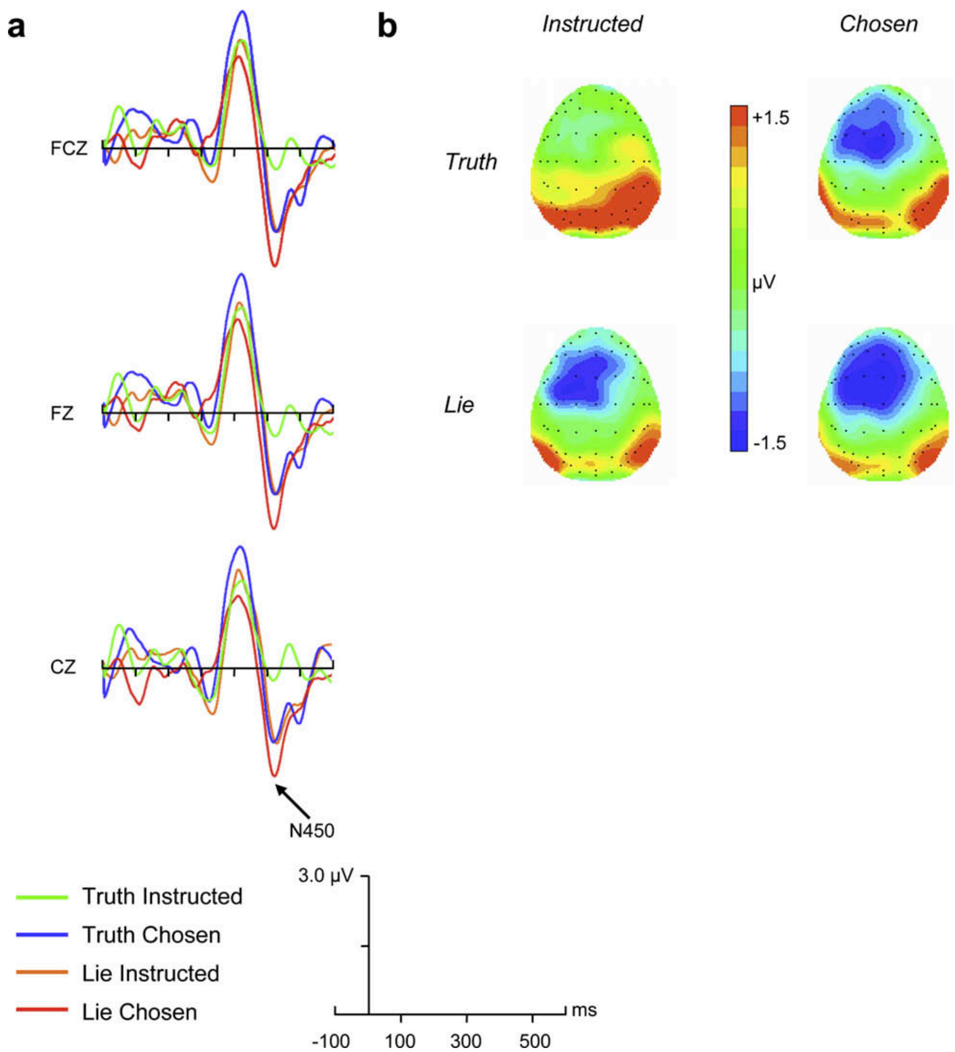

Fig. 2. (a) Grand averaged ERPs elicited by the truth instructed, lie instructed, truth chosen, and lie chosen trials from frontal–central (FCZ), frontal (FZ), and central (CZ) electrode sites. Lie instructed and lie chosen trials give rise to an enhanced negativity starting at 400 ms. ERPs associated with truth chosen trials show a similar late negativity. (b) Potential maps depicting the scalp distribution for the observed MFN (400–500 ms). The topography of the negativity elicited by the lie instructed, lie chosen, and truth chosen (deceptive truth) trials was maximum at fronto–central electrode sites, pointing to the anterior cingulate cortex (ACC) as the neural site of conflict processing.

their MFN during the chosen truth trials, $r = -.47$, $p < .001$. This suggests that the better participants' mentalizing ability, the more conflict they experienced while telling deceptive truths. The MFN amplitude during the other three trial types was not significantly correlated with the MIE score (chosen lie: $r = -.19$; instructed truth: $r = -.37$; and instructed lie: $r = -.27$). Furthermore, we created a measure of deceptive ability by analyzing how many times a participant successfully deceived their opponent during the game. There was no significant difference in the overall deception success for chosen truth and chosen lie trials ($p = .64$), indicating that the opponent found it equally difficult to detect lies and deceptive truths. Rather surprisingly, however, MIE scores were negatively correlated with success at deceiving the opponent during the choice condition, $r = -.65$, $p < .001$. Thus, the better participants' mentalizing ability, the less successfully they performed the deception task. This indicates that mentalizing skills may actually have interfered with deception performance.

5. Discussion

Although previous findings suggest that increased demands on cognitive control arise selectively when individuals make counterfactual statements, the current findings demonstrate that deceptive intent is critical. We found clear evidence for increased recruitment of cognitive control processes when participants expressed the truth in an attempt to deceive an interaction partner compared to when they expressed the truth without deceptive intention. Indeed, the ERP correlate of cognitive control elicited by the chosen truth trials was not only similar to the correlate elicited by lie trials, but also closely resembles the cognitive control correlates identified in previous studies (West et al., 2005, 2004). Thus, the recruitment of cognitive control processes during deception cannot be reduced to the need to inhibit a tendency to state the truth. Rather, the challenge for the liar is likely in keeping their own and another's mental representation of the current situation apart.

The finding that the higher participants scored on a mentalizing task, the more conflict the ERP correlate showed when telling the truth deceptively, is in line with this interpretation. It seems plausible that a high score in the MIE test indicates not only that an individual can accurately ascribe mental states to others, but rather that this individual is more prone than others to do so. Thus, individuals who thought about the opponent's mental state while telling the truth deceptively may have experienced more conflict than those individuals who thought less about the other. Interestingly, we also found evidence for a link between mentalizing skills and deceptive ability. The better participants performed in the mentalizing task, the worse they were at deceiving their opponent. At first, this result may seem counterintuitive. Surely, without a theory of mind deception would be unattainable (Hala & Russell, 2001). A possible explanation is that participants with better mentalizing skills experienced more conflict, which interfered with their ability to deceive the other. Moreover, one could speculate that individuals with high mentalizing skills are more empathic, making it difficult for them to deceive another person, and to control their own facial expression during deception. We hope that future research will shed light on this interesting issue.

Taken together, our results suggest that we can gain a better understanding of the cognitive and neural processes underlying deception by studying deception in social context, and by focusing on people's intentions. Rather than classifying statements as true or false, it may be more fruitful in the search for cognitive and neural markers of deception to consider the intention driving true or false statements. Of course, poet William Blake knew this all along, stating in the Auguries of Innocence "A truth that's told with bad intent beats all the lies you can invent".

Acknowledgements

This study was conducted at Rutgers University, NJ, with support of the School of Psychology, Newark to N.S. We thank Ekaterina Dobryakova for her help with data collection. We would also like to thank two anonymous reviewers for their comments.

References

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

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). *Journal of Child Psychology and Psychiatry, 42*, 241–252.

Blakemore, S. J., Winston, J., & Fruth, U. (2004). Social cognitive neuroscience: Where are we heading? *Trends in Cognitive Sciences, 8*, 216–222.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108*, 624–652.

Dennett, D. C. (1987). *The intentional stance*. Cambridge: MIT Press.

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society B: Biological Sciences, 358*, 459–473.

Ganis, S. M., Kosslyn, S. M., Stose, S., Thompson, W. L., & Yurgelun-Todd, D. A. (2003). Neural correlates of different types of deception: An fMRI investigation. *Cerebral Cortex, 13*, 830–836.

Gratton, G., Coles, M., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology, 55*, 468–484.

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24*, 95–112.

Hala, S., & Russell, J. (2001). Executive control within strategic deception: A window on early cognitive development? *Journal of Experimental Child Psychology, 80*, 112–141.

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, 42*, 878–901.

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.

Kozel, F. A., Padgett, T. M., & George, M. S. (2004). A replication study of the neural correlates of deception. *Behavioral Neuroscience, 118*, 852–856.

McCarthy, G., & Wood, C. C. (1985). Scalp distributions of event-related potentials: An ambiguity associated with analysis of variance models. *Electroencephalography and Clinical Neurophysiology, 62*, 203–208.

Nuñez, J. M., Casey, B. J., Egner, T., Hare, T., Hirsch, J., et al. (2005). Intentional false responding shares neural substrates with response conflict and cognitive control. *Neuroimage, 25*, 267–277.

Proust, J. (2007). Metacognition and metarepresentation: Is a self-directed theory of mind a precondition for metacognition? *Synthese, 159*, 271–295.

Sip, K. E., Roepstorff, A., McGregor, W., & Frith, C. D. (2008). Detecting deception: The scope and limits. *Trends in Cognitive Science, 12*, 48–53.

Spence, S. A. et al. (2004). A cognitive neurobiological account of deception: Evidence from functional neuroimaging. *Philosophical Transactions of the Royal Society B: Biological Sciences, 359*, 1755–1762.

Spence, S. A., Farrow, T. F., Herford, A. E., Wilkinson, I. D., Zheng, Y., & Woodruff, P. W. (2001). Behavioural and functional anatomical correlates of deception in humans. *Neuroreport, 12*, 2849–2853.

West, R., Jakubek, K., Wymbs, N., Perry, M., & Moore, K. (2005). Neural correlates of conflict processing. *Experimental Brain Research, 167*, 38–48.

West, R., Bowry, R., & McConville, C. (2004). Sensitivity of medial frontal cortex to response and nonresponse conflict. *Psychophysiology, 41*, 739–748.