

The production and detection of deception in an interactive game

Kamila E. Sip^{a,b,*}, Morten Lynge^d, Mikkel Wallentin^a, William B. McGregor^c, Christopher D. Frith^{a,e}, Andreas Roepstorff^{a,c}

^a Center of Functionally Integrative Neuroscience (CFIN), Aarhus University Hospital, 8000 Aarhus, Denmark

^b Department of Psychology, Rutgers University, Newark, United States

^c Institute of Anthropology, Archaeology and Linguistics, University of Aarhus, Denmark

^d Department of Computer Science, University of Aarhus, Denmark

^e Wellcome Trust Center for Neuroimaging, Functional Imaging Laboratory, University College London, United Kingdom

ARTICLE INFO

Article history:

Received 2 March 2010

Received in revised form 11 August 2010

Accepted 12 August 2010

Available online 19 August 2010

Keywords:

Deception

Social interaction

Decision-making

ABSTRACT

This experiment tests how people produce and detect deception while playing a computerized version of the dice game, Meyer. Deception is an integral part of this game, and the participants played it as in real life, without constraints on whether or when to attempt to deceive their opponent, and whether or when to accuse them of deception. We stress that deception is a complex act that cannot be exclusively associated with telling a falsehood, and that it is facilitated by hierarchical decision-making and risk evaluation. In comparison with a non-competitive control condition, both claiming truthfully and claiming falsely were associated with activity in fronto-polar cortex (BA10). However, relative to true claims, false claims were associated with greater activity in the premotor and parietal cortices. We speculate that the activity in BA10 is associated with the development of high-level executive strategies involved in both types of claim, while the premotor and parietal activity is associated with the need to select which particular claim to make.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

The question of how people deceive and how they detect deception has been widely discussed in the literature of psychology, forensics, and recently, neuroimaging (Ganis, Kosslyn, Stose, Thompson, & Yurgelun-Todd, 2003; Kozel et al., 2005; Langleben et al., 2005; Spence, Kaylor-Hughes, Farrow, & Wilkinson, 2008). As is widely recognized (e.g. Greely & Illes, 2007; Sip, Roepstorff, McGregor, & Frith, 2008a; Sip, Roepstorff, McGregor, & Frith, 2008b; see also Abe, Suzuki, Mori, Itoh, & Fujii, 2007; Greene & Paxton, 2009), one of the main questions for experimental studies of deception is the degree to which laboratory tasks can elicit the same processes as those occurring in 'real life' settings. However, the closer we come to a real life setting in the laboratory, the more we must sacrifice tight experimental control.

In one of the first studies attempting to address the issue of 'instructed lies', Abe and colleagues (2007) introduced a very elegant twist to their experimental protocol. The twist takes place at the level of the instructions, where the participants received two mutually contradictory instructions from two experimenters. In

practice, during the break necessary for the radioactive tracer to decay, when experimenter 1 was not present in the room, experimenter 2 secretly asked the participants to deceive experimenter 1 by answering questions with opposite responses than those suggested by the experimenter 1. The main effect of deceiving one of the experimenters, that is 'falsifying' otherwise truthful responses, showed increased neural activation of the left dorsolateral and right anterior prefrontal cortices. This study, though it does resonate with real-life deception, still suffers from a problematic issue of 'instructed lies'. In this scenario, the participants faced an externally introduced change to suddenly obey a different set of rules. It is difficult then to account for that change in terms of neural activation in reference to the processing of deception as the actual peripheral attentional activation and deception activation could have contributed to the final results. Therefore, to address these issues in our investigation and to avoid instructing the participants to 'lie' at any stage (see also Baumgartner, Fischbacher, Feierabend, Lutz, & Fehr, 2009; Greene & Paxton, 2009), we chose a paradigm which fully allows the participants to be deceptive at will.

As a compromise between ecological validity and experimental control, we have developed a laboratory version of the deception game 'Meyer', widely played in Denmark, in a form suitable for the scanning environment. One advantage of using this game is that participants must decide for themselves when to deceive, a decision based on the current state of the game (see below for a detailed account of this game). This is in contrast to the many studies of

* Corresponding author at: Department of Psychology, Rutgers University, 101 Warren Street, Smith Hall 301, Newark 07102, United States. Tel.: +1 973 353 1827; fax: +1 973 353 1171.

E-mail addresses: ksip@psychology.rutgers.edu, kami.sip@gmail.com (K.E. Sip).

deception in which participants are instructed by the experimenter exactly when to be deceptive, that is to tell a falsehood (see e.g. Ganis et al., 2003; Langleben et al., 2002, 2005; Mohamed et al., 2006). Another advantage is, that, by using a two-person game, we can look at the detection as well as the production of deception. The corresponding disadvantages with the game include loss of control over how many deceptive events will be generated and, in the detection condition, the participants' expectation that deception will occur.

Our use of a two-person game emphasizes the inherent communicative nature of deception. We characterize deception as an intentional attempt conducted by an agent (A) to induce or reinforce in another individual (recipient) a belief that, in A's opinion, is false (Sip, 2009). One implication of this characterization is that deception is interpersonal and interactive; another is that being deceptive is not synonymous with telling a falsehood (see also Abe et al., 2007; Greene & Paxton, 2009). The agent may intentionally create the belief in the recipient that he is 'lying' when he is actually telling the truth. Such intentions are an important feature in the game of Meyer in which two players take it in turns to throw two dice. The players may attempt to accomplish deception by telling the truth with a deceptive intention. The players are effectively exchanging statements, referred to in this paper as claims. Players in turn make claims about the state of the dice (production moves), and respond with either acceptance or rejection of the other player's claim (responding moves). The game progresses much in the manner of a conversation, with interlocked turns, or moves. A player's claims are not isolated responses, but are tied tightly into the social context and history of the game in which they occur. Players must take the intentions of the claimant into account, requiring inferencing informed by the interpersonal and interactive nature of the game.

Although many different paradigms have been used in neuroimaging studies of deception (for reviews see Greely & Illes, 2007; Sip et al., 2008a), there is a consistency in the pattern of brain regions activated (Christ, Van Essen, Watson, Brubaker, & McDermott, 2009): anterior cingulate (ACC), dorsolateral prefrontal cortex (DLPFC), medial prefrontal cortex (MPFC), and inferior frontal gyrus (IFG). However, these regions are clearly not dedicated solely to deceptive behaviour. Rather, deception draws on a number of cognitive processes that are not uniquely concerned with deception. For successful deception, there is a need to plan ahead, to inhibit prepotent responses (i.e. to tell the truth), and to keep track of the opponent's beliefs, all of which processes are associated with activity in the prefrontal cortex (e.g. Amodio & Frith, 2006; Koechlin & Summerfield, 2007; Shallice, 1988).

Recently, Koechlin and colleagues (Koechlin, Ody, & Kouneiher, 2003; Koechlin & Summerfield, 2007; Koechlin & Hyafil, 2007, for review see Ramnani & Owen, 2004) have developed a hierarchical model of cognitive control of executive brain functions in action selection. In the latest version (Koechlin & Summerfield, 2007), the model effectively identifies four levels of cognitive control of action with respect to earlier events. The model defines the process of action selection as a "fractionation of cognitive control according to the temporal framing of actions and events involved in selection" (Koechlin & Summerfield, 2007, see also Braver & Bongiolatti, 2002). The first and lowest level is a *sensory control* in selecting motor action, with activity in premotor regions. The second level is associated with a *contextual control* involved in responding to stimuli in the immediate external context in which these occur with activity located in caudal LPFC (typically BA9/44/45). The third level is an *episodic control* with specific activity typically in DLPFC (BA46). Finally, the fourth level, a subdivision of the episodic control, is a *branching control* located in the lateral fronto-polar cortex (BA10).

We would expect these levels to be present in decision-making in Meyer, however, in contrast with many other paradigms aimed

at studying deception, the most prominent aspect of action selection in this study lies in the branching control. This is due to the fact that the episodic control is assumed to be divided into a control of signals, both subsequent and preceding ones (Koechlin & Summerfield, 2007). As the branching control appears to allow for accommodating a pending choice to an ongoing behaviour, it significantly contributes to the construction of a coherent deceptive context which accounts for past and upcoming events. In the game of Meyer it is of particular importance for the player to decide whether or not to be deceptive after each dice throw in relation to past choices and the stage of the game.

There is a difference between true and false claims, specific to Meyer, which is not directly related to deception. When making a true claim the player reports the value of the dice throw. When making a false claim, the player must choose which claim to make. Thus, when making a false claim, additional processes will be engaged concerned with selecting responses from among competing alternatives. This process is likely to involve a frontal-parietal network (see Schumacher, Cole, & D'Esposito, 2007 for a review).

Interestingly, much less previous work was done on detecting deception (responding moves in Meyer) (see e.g. Grezes, Frith, & Passingham, 2004). In Meyer, the player has to decide whether the opponent is being truthful. If the player accepts the opponents claim there is no immediate consequence, since the game continues. However, challenging the opponent's claim is a more risky strategy with an uncertain outcome, since the player may lose.

2. Materials and methods

2.1. Stimuli

Participants were examined in a 3 T scanner while playing a computerized version of Meyer with a human opponent outside of the scanner (confederate). The participant and the opponent played in turn, as in the standard game of Meyer. However, the data collected during the times that opponent was making his choices were not included in the analysis. The game was divided into 4 game sessions of 5 min each and one 5 min control session. The game sessions involved each of the 4 possible combinations of the 2 variables: (A) the face of the opponent was displayed/not displayed to the subject live on screen (see Fig. 1); and (B) subject was informed that their face was displayed/not displayed to their opponent via a small Internet camera (which was in fact merely the outer shell of an Internet camera). The order of all sessions was randomized within and across subjects. Surprisingly, no effects of this manipulation were evident in the imaging data, and so the conditions are collapsed below.

In this game, the values of the throws increase in the following order: 32, 41, 42, 43, 51, 52, 53, 54, 11, 22, 33, 44, 55, 31 (Little Meyer), and 21 (Meyer). The order of dice combinations presented on the screen was randomized (see Fig. 1). For play to continue, a player must claim at least as high or higher dice combination than that claimed by the opponent on the previous trial, e.g. if the opponents' dice combination was 44 on the previous trial, then the participant needs to claim at least 44 (see Fig. 2). If this claim is false and the player is challenged by the opponent, then the player loses and a new round begins. Otherwise the game continues with a next dice throw. However, if the claim is true and the player is challenged by the opponent, then the opponent loses. Thus, it can be an advantage to the player to deceive the opponent into falsely believing that the player is making a false claim. In this game the difference between true and false claims is not so clear-cut in terms of deceptive intent.

The control session involved the same procedure as the four game sessions in most respects except that deception was not involved and the participants could not see the face of their opponent. On production trials the player claimed their dice throw by choosing the identical combination from the configurations shown on the screen regardless its value and with no consequences to the progress of exchange. On the responding trials, the players indicated BELIEVE to all claims regardless its value. No money was at stake. This control, however, does not differ from the game sessions in terms of attention and cognitive complexity in matching the dice combinations.

After the rules of the game were explained, the participants spent a few minutes playing the game outside the scanner to ensure that they understood the rules and to get familiarize with the interface.

The participants interfaced with the game via a screen (for details see Fig. 1). Each production and responding trial was separated by an instrumental interval of 2.5 s (Dale & Buckner, 1997) to allow partial recovery of the hemodynamic response.

Participants were paid 100 Danish kroner (€ 15) for participation. In addition each subject began with a kitty of another 100 Danish kroner (€ 15) to play with. Each point lost as a result of being detected making a false claim or wrongly accusing the opponent of making a false claim, lead to reduction in the subject's pool of money

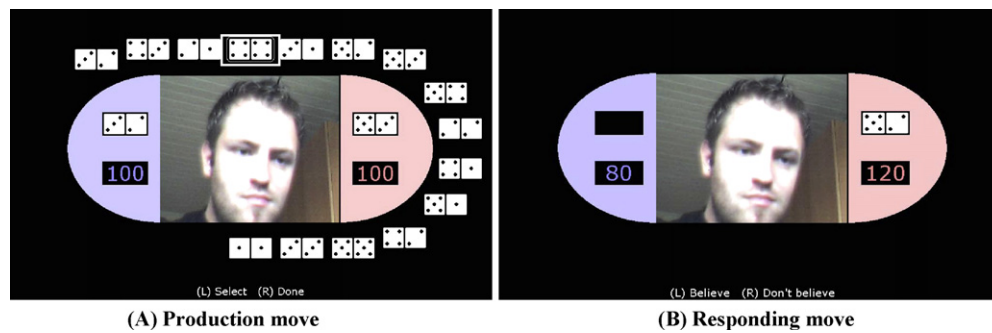


Fig. 1. Interface examples of move types during the computerized Meyer game. (A) The left side of the screen illustrates a subject's current points and the computer-generated dice throw. The right side shows the opponent's points and his previous claim. All possible dice combinations are shown in a randomized order to allow the subject to select their claim and make an active decision. Just above the opponent face (depicted) one combination is enclosed within a frame; this is a selectable combination. The dice combination in the frame at the beginning of the selection procedure was chosen at random. To make their choice of dice combination to claim, the subjects click the left button on the button pad to move successively one combination at a time counter-clockwise until the desired combination falls within the frame above the picture of the opponent. They then make their claim by clicking the right button. (B) Again, the left side of the screen shows a subject's current points. The right side indicates the opponent's points, and above is shown his last claim. Here, he needs to decide whether to accept ("Believe") or reject ("Don't believe") the claim. This is done with, respectively, a left or a right button press.

and an equivalent increase in the opponent's pool. Every point gained as a result of catching the opponent claiming falsely or being wrongly accused of a false claim lead to an increase in the subject's financial balance and an equivalent reduction in the opponent's. This means that per round, the participants could lose or gain 10 Danish kroner for any combination other than '21' Meyer, and 20 Danish kroner for Meyer. In the entire game, the participants could maximally triple their kitty money, or lose it entirely in the game. They could not lose their participation fee.

After scanning, participants were interviewed concerning their feelings about the game and the strategies and cognitive processes they employed. These post-scanning interviews aimed at obtaining more detailed information on the subject's decision-making processes, emotions, and attitudes during the game.

2.2. Participants

18 right-handed participants with no reported neurological or psychological disorders took part in the experiment, 8 females. Four participants (3 females) were discarded due to technical problems. The participants were local students (6) or international exchange students at the University of Aarhus and Aarhus Business School, Denmark, right-handed, with a mean age of 24 ± 1.33 . All participants gave their informed written consent to take part in the experiment. Ethical approval was

granted by Den Videnskabssetiske Komite for Aarhus AMT, number 20050206. The international students were all novices at the game, whereas the local students were somehow familiar with the game, yet claimed neither to have played it for a long time nor be proficient at it.

2.3. fMRI image acquisition

Functional images were acquired on a GE 3 Tesla system (Signa Excite, General Electric, Milwaukee, WI) with a standard head coil (8 channel). T1 structural and T2*-weighted images were obtained using an echo planar imaging sequence with the following parameters: repetition time (TR): 3000 ms, echo time (TE): 30 ms, flip angle: 90. 35 sequential, descending 3.5 mm axial slices (no gap between slices) were obtained per volume of 100 in each session, with an in plane resolution of $3 \text{ mm} \times 3 \text{ mm}$. The voxels were interpolated into $2 \text{ mm} \times 2 \text{ mm} \times 2 \text{ mm}$ during pre-processing.

2.4. fMRI data analysis

Raw images were reconstructed and converted for further analysis, spatially realigned and motion-corrected according to all six motion parameters, slice time

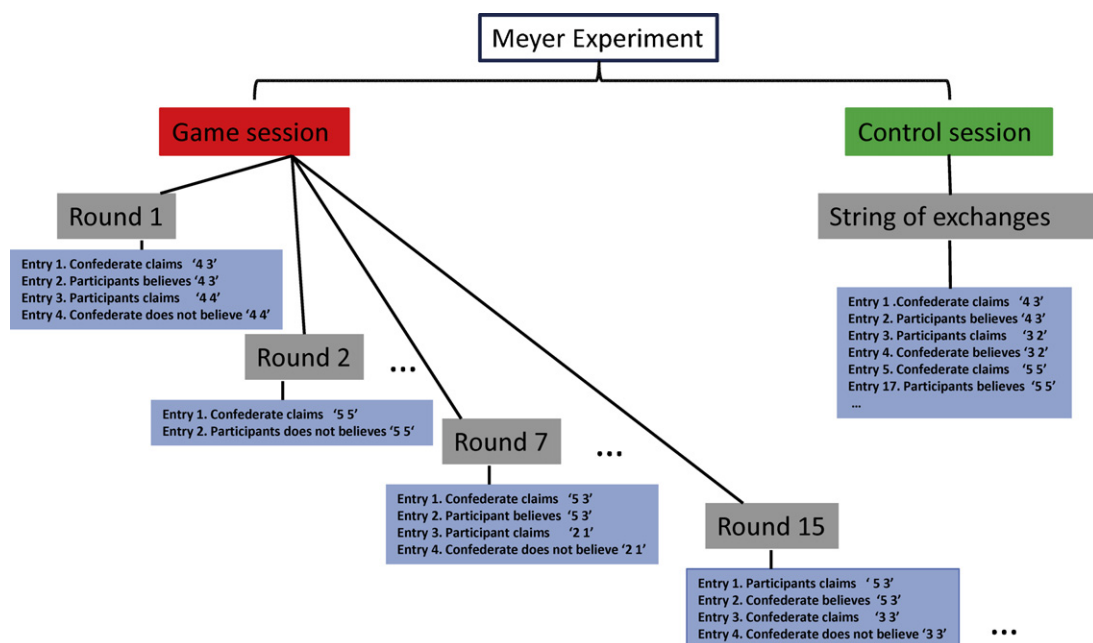


Fig. 2. The progress of Meyer game in game and control sessions. Each experimental session lasted 300 s. In the control session, the participants claimed the exact dice throw they got from the computer and believed every claim of their opponent's. In the game sessions, participants had free choice as to whether believe claims of the opponent and whether to be deceptive. In game sessions, the number of rounds and their length differs within and across participants. For example, for participant XX, Round 1 ends after four entries with confederate disbelieving a true claim. This means that he loses points and needs to start the next round. Round 2 ends after only two entries with a participant disbelieving a true claim. Round 7 ends with the confederate correctly disbelieving a false claim. Game continues.

corrected, and normalized to the MNI template using SPM5 (Statistical Parametric Mapping, Wellcome Trust Centre for Neuroimaging, London, UK) and executed in MATLAB (Mathworks Inc., Sherborn, MA, USA). After normalization, the images were smoothed with an 8-mm full-width at half-maximum Gaussian filter.

A design matrix was constructed, including 6 regressors: claiming truthfully, claiming falsely, accepting opponent's claim, rejecting opponent's claim, and the two regressors from the control session: production and responding. All events were modelled using the standard hemodynamic response function of SPM5. This model has proven a reasonable and comprehensive model of the hemodynamic response function (Henson & Friston, 2007). The fit to the data was estimated for each participant using a general linear model (Friston et al., 1995) with a 128 s high-pass filter, global scaling and AR(1) modelling of serial correlations.

The reported effects were modelled in terms of both decision-making and response selection: for production the event duration was from the moment the participants saw their computer-generated dice throw until the moment they made their claim; for responding, events lasted from the time the opponent's claim was displayed until the moment they accepted or rejected it.

Individual *T*-contrasts related to the different task effects were created from the estimated weights, and these were used in a second level random effects analysis (a one-tailed *t*-test for each first level contrast type) in order to facilitate inferences about population effects.

Whole brain results are reported with a threshold of $p < 0.05$, FDR corrected for multiple comparisons at voxel level. Additionally, we indicate several areas which survive a more stringent FWE correction. Given that the complete analysis of responding moves does not survive the corrected threshold, for the sake of completeness we conducted an exploratory analysis on the affected contrasts. The threshold there was at $p < 0.001$, uncorrected for multiple comparisons and a cluster-size threshold of 40 voxels (see Supplementary Material).

3. Behavioural results and reports

Since the experiment employed a real-life paradigm, neither the number of true and false trials nor their placement was controlled. Games consisted of an average of 35.93 ± 0.97 production moves and 34.43 ± 0.56 responding moves. On average, in production moves participants claimed truthfully 21.1 ± 1.1 times, and falsely 14.1 ± 0.9 times. In responding moves they rejected their opponent's claim on average 15.6 ± 0.34 times, and accepted it 18.9 ± 0.6 times. In the control session, participants claimed truthfully 6.79 ± 0.23 times, and accepted the opponent's claims 9.93 ± 1.23 times.

On average participants won on about half the trials ($48 \pm 2.7\%$ SEM). They claimed falsely on just below half of their production moves ($40 \pm 2.7\%$) (see above). They were wrongly accused of claiming falsely on about half of their production moves ($52 \pm 2.9\%$), and were correctly accused on about a quarter of them ($23 \pm 1.38\%$). Participants wrongly accused their opponent of claiming falsely on slightly more than half of the responding moves in which they rejected the opponent's claim ($63 \pm 2.7\%$), and correctly accused them in about a third ($37 \pm 2.7\%$). These statistics underline the riskiness of responding moves of accusation: in a little over half of this type of move the claimant is wrong, and thus loses. They also reveal that our study was relatively well balanced in terms of number of events for each condition, even though we did not attempt to control this variable.

Reaction times for production moves were much longer than for responding moves. This is presumably at least partly due to the fact that participants had to select from among the 15 dice configurations displayed on screen, rather than between just YES and NO buttons. For responding moves, participants responded more quickly when accepting the claim (pressing YES) than when challenging the claim (pressing NO) (on average 3.5 ± 0.23 s with a minimum of 2.41 s and max of 5.7 s; 4.16 ± 0.30 s with a minimum of 2 s and maximum of 6.01 s; two-tailed pair-wise *t*-test, $p < 0.04$, $t(13) = 2.23$) (see Fig. 3). This is in keeping with the above discussion on the risk associated with challenging claims.

Production moves, however, showed no significant differences in reaction times for false (9.69 ± 0.92 s with a minimum of 5.88 s and a maximum of 13.20 s) vs. true claims (9.53 ± 0.52 s with a minimum of 7.6 s and a maximum of 13 s). These very long times, as just

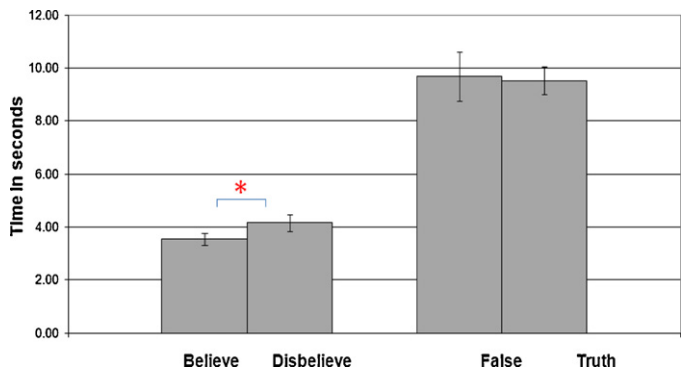


Fig. 3. Response time during decision-making. This figure represents the reaction times associated with two types of deception, production (False, Truth) and detection (Accept, Reject). There is a significant difference between reaction times in the responding moves ($p < 0.04$, $t(13) = 2.23$). The difference between claiming true or false is not significant $p < 0.81$, $t(13) = 0.24$.

mentioned, reflect the time required for selecting among the configurations shown on screen, rather than processing time. However, this is only part of the story. The post-scanning reports confirmed that some participants sometimes deliberately prolonged their decision-making time when claiming truthfully in the hope that this would cause their opponent to disbelieve them, and induce them to challenge the claim. This strategy presumably accounts for why claiming falsely did not take significantly longer than claiming truthfully.

In the control session, to state the truth the average time was 8.05 ± 0.69 s with a minimum time of 4.1 s, and a maximum time of 11.43 s. The time to accept the claim was 2.40 ± 0.18 s on average with a minimum of 1.5 s, and a maximum of 3.4 s.

4. Imaging results: production moves

As already mentioned, in the control session, in production moves the subject simply reported the current throw of the dice, whereas in production moves in the game sessions participants could claim either falsely or truthfully. In contrast to the control condition, claiming falsely in the game session was associated with activity in a network of frontal and parietal regions (see Table 1 and Fig. 4).

In comparison to the control condition, claiming truthfully activated a subset of these regions, in particular the prefrontal region, BA10 (Table 1 and Fig. 4). When claiming falsely is directly contrasted with claiming truthfully, activity is seen in the premotor and parietal cortices. In agreement with other investigations of deception, there were no areas where claiming truthfully elicited more activity than claiming falsely.

5. Imaging results: responding moves

Surprisingly, the comparison between rejecting and accepting the opponent's claim does not survive the corrected threshold (see Supplementary Materials for Details). Unexpectedly, there was a significant decrease of activity in the posterior STS when claims were accepted or rejected in game sessions in comparison to control sessions (see Fig. 5).

6. Discussion

In the current study, we presented a deceptive dice game in which players expect their opponents to attempt to deceive them, and are perhaps more alert to the possibility of deception than in casual conversation. This expectation is exploited by players, who

Table 1
Brain regions showing activation in production moves.

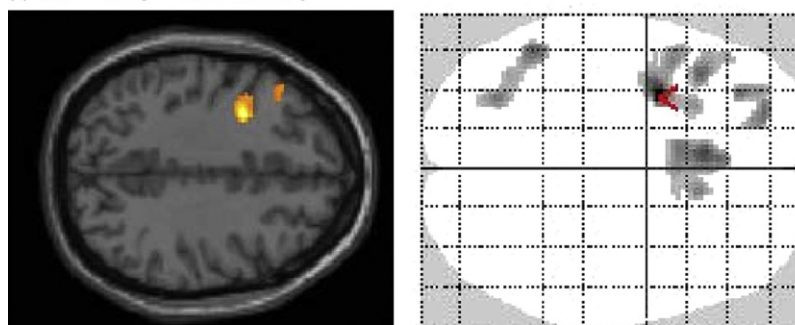
Brain region	Cluster size	x	y	z	z-value	T-value	FDR p-voxel	FWE p-voxel
Claim falsely – control								
Left premotor, BA6	20	–24	12	56	5.49	11.35	<0.001	<0.01
Left prefrontal, BA10	6	–28	54	14	5.22	10.01	<0.001	<0.02
SMA, BA6	1	–10	16	58	5.13	9.62	<0.001	<0.03
Left SPL, BA7	11	–42	–56	48	5.33	10.55	<0.001	<0.01
Precuneus, BA7	8	0	–66	48	5.20	9.89	<0.001	<0.02
Claim truthfully – control								
Left prefrontal, BA10	767 (16)	–44	48	10	5.49	11.36	<0.001	<0.003
Left prefrontal, BA10	Subsidiary peak within same cluster	–28	50	20	4.00	5.79	<0.02	ns
Left prefrontal, BA10	Subsidiary peak within same cluster	–32	52	8	4.92	8.70	<0.004	ns
Right prefrontal, BA10	95	20	64	8	4.21	6.35	<0.02	ns
Claim falsely – claim truthfully								
Left premotor, BA6	18	–30	6	38	5.08	9.36	<0.005	<0.03

The coordinates are given according to the MNI space, together with its *T*-scores, *Z*-scores and significant thresholds $p < 0.05$ (all FDR corrected for multiple comparisons). Additionally, we indicate the areas which survive more stringent threshold of FWE correction (ns – non-significant at FWE level).

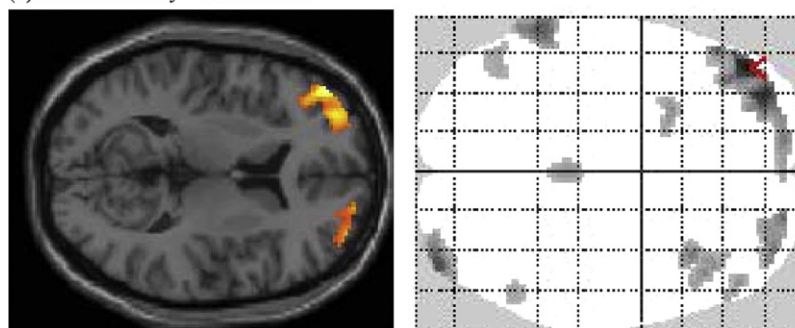
make true claims, but do so in ways designed to invoke the inference that they are making a false one. In the context of Meyer, a true claim is often used with as much deceptive intent as a false one. An advantage of this strategy is that if a player can persuade

their opponent that they are falsely claiming above the value of their throw when they are actually claiming truthfully, and consequently invite the other to challenge them, they will win. This kind of complex decision-making is a common practice in everyday

(a) Claim falsely – claim truthfully



(b) Claim truthfully – control truth



(c) Claim falsely – control truth

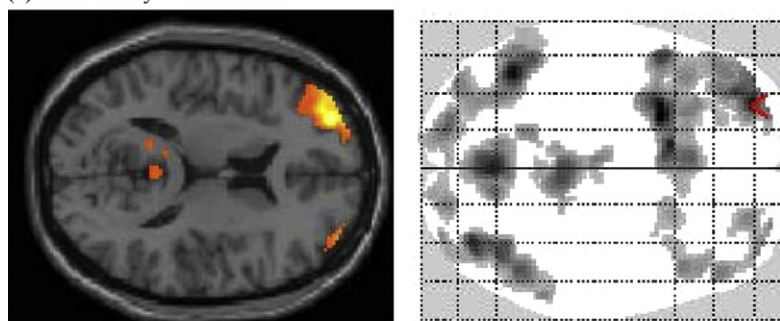


Fig. 4. View of glass brain from above in production moves. This figure represents the main peak activations in the three contrasts. For illustration purposes these contrasts are presented at the uncorrected level ($p < 0.001$). The arrows indicate global maxima.

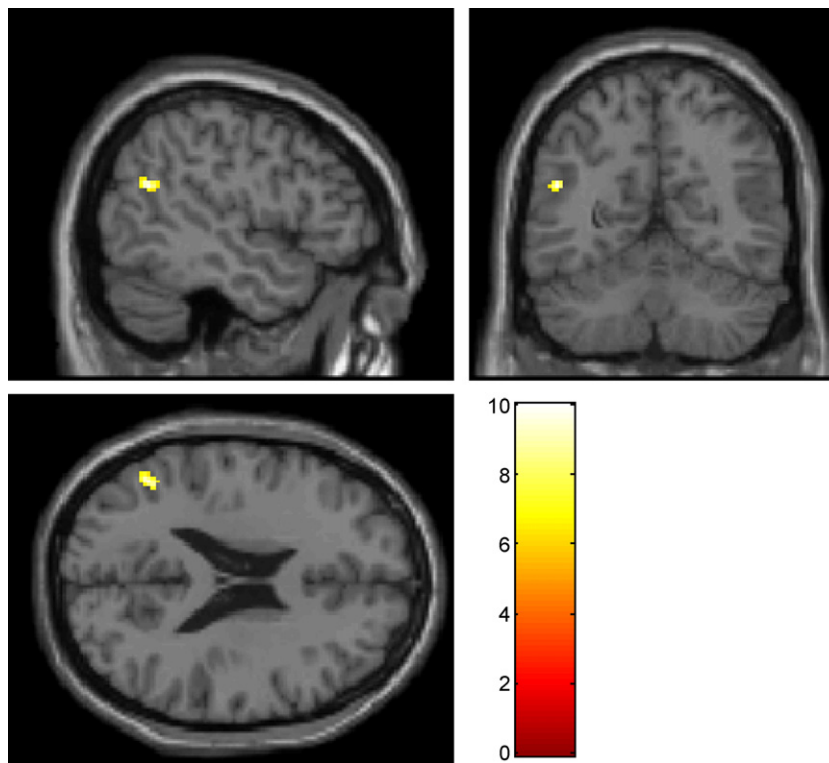


Fig. 5. Decreased activation associated with responding moves in game sessions. A decreased activity is seen in left posterior STS ($-48, -58, 22$, T -value 10, z -value 5.22, $p < 0.05$ FDR corrected) when participants decide whether to believe or disbelieve the opponent's claims. The colour bar corresponds to T -values.

deception indicating that it may be too narrow to analyze deception only in terms of false claims.

Most previous studies have found activity in the prefrontal cortex to be related to deception. However, this activity has usually been located in DLPFC (BA46/9) and ACC (see for review Sip et al., 2008a; also Spence et al., 2004), rather than the frontopolar cortex (BA10). According to Koechlin's framework (Koechlin & Summerfield, 2007; Koechlin & Hyafil, 2007), DLPFC is lower in the executive hierarchy than BA10 and is concerned with selection of responses appropriate in the current context (Frith, 2000; see also Miller & Cohen, 2001). This is consistent with the observation that in most studies of deception participants are instructed when to report falsely, when truthfully. We do not observe the activation of DLPFC, but then our task did not involve decision-making at the intermediate levels. In most other experiments, free choice is absent, and the decision as to whether or when to deceive is pre-established by the experimenter and thus 'intermediate' areas are invoked (Roepstorff & Frith, 2004).

A few recent neuroimaging experiments have investigated deception in more ecologically valid paradigms (see e.g. Baumgartner et al., 2009; Greene & Paxton, 2009; Spence et al., 2008). In a study by Spence et al. (2008), the researchers tested the effect of admitting to intimate and potentially embarrassing autobiographical material. Although this experiment goes some way towards resolving the problems of deception-on-cue experiments, it introduces a different problem, the effects of which are difficult to ascertain. The subjects were instructed to attempt to maintain an approximate balance between truthful and untruthful responses, which they appear to have done quite successfully. Not only is this likely to represent a much higher frequency of untruthful utterances than in everyday interaction, but more importantly, it is impossible to know what neurological effects this instruction may have had, or by what strategies subjects achieved the approximate balance.

In a study conducted by Baumgartner et al. (2009), the researchers allowed participants to decide whether to be honest or dishonest in a well established trust game. They introduced an additional element to the game by asking the participants to make a promise about their subsequent behaviour and allowing them to either keep or break the previously expressed promise. This paradigm resembles ours in two ways: (1) in both studies the participants tried to deceive their partners and benefit from their loss, and (2) the participants made active decisions to behave dishonesty towards another person in a real-life interaction. The results revealed that, when breaking their promise, the players showed greater activity in the anterior cingulate cortex, parts of the dorsal prefrontal cortex, and the amygdala. Interestingly, already in the decision stage to be dishonest the authors observed activation in left BA10/46 as in our study.

Our results are also consistent with findings reported recently by another group. Greene and Paxton (2009) investigated two competing hypotheses regarding the cognitive and neural nature of honesty: (1) honest choices result from active resistance of temptation to tell falsehood, and (2) honest choices result from the absence of temptation (Greene & Paxton, 2009). The participants were asked to predict an outcome of a coin-flip, and were then either punished or rewarded with money for their accuracy. The authors observed bilateral prefrontal activity in DLPFC associated with making true responses by dishonest participants (Greene & Paxton, 2009), which seem to overlap conceptually and functionally with telling the truth with a potential deceptive intention as exercised in our study. Interestingly, the values of punishment and reward did not seem to be a factor in the decision to 'lie' and thus this result does not seem to be driven by the monetary reward (Greene & Paxton, 2009: 12508). This may suggest that in experimental deception, even with monetary incentives as in both our investigation and Greene and Paxton's, the free choice to exercise deception seems to be a much bigger factor in

decision-making with respect to consequences and responsibility for actions.

6.1. Production moves

In our study, both in the control condition and when claiming truthfully during game sessions, players were effectively reporting the dice configurations they had thrown. Although the semantics and the truth value of these claims are identical, the context of the game is very different. This difference is also seen in the functional imaging data, where the truthful claims in the game sessions, as distinct from the control sessions, elicit activity in a large network in the prefrontal cortex, typically associated with executive function and decision-making (e.g. Fig. 4).

In comparison to the control task, increased activity is seen in the prefrontal cortex when claiming falsely. However, this increased activity is otherwise very similar to the pattern of activity seen when claiming truthfully (compare Fig. 4b and c). Both claim types were associated with activity in the fronto-polar cortex, BA10. This region of the prefrontal cortex is considered to be at the top of a hierarchy of executive functions instantiated in the frontal cortex. These play an important role in tasks involving actions that are contingent upon unpredictable future events (Koechlin et al., 2003; Koechlin & Summerfield, 2007; Koechlin & Hyafil, 2007).

The function of BA10 is extremely broad in cognitive paradigms. Burgess and colleagues (2007) discuss the gateway hypothesis of area BA10 in which they indicate that this area may be activated in various paradigms ranging from simple to most complex cases of judgment, problem-solving and perception involving memory retrieval (Burgess, Dumontheil, & Gilbert, 2007). The fronto-polar cortex, however, plays a crucial role in decision-making where the expected behaviour is not adequately specified and thus there are many course of actions that may be taken. In relation to our paradigm this function seems to be well captured when the participants make choices in the context of the game.

An increased activation in left BA10 (in addition to ACC, VLPFC and MPFC) was also observed in Abe et al. (2006) and associated with both types of deception (pretending to know and pretending not to know). The activation of BA10 with its peak in $-28, 54$, and 14 is consistent with our finding. However, it is important to note that BA10 was but one among many brain areas that have been consistently reported in 'instructed lies' paradigms. It is thus difficult to link these results with those reported in current study given the substantial differences in experimental design.

The speculations about the activation of the fronto-polar cortex observed in current study differ from that found by Ganis and colleagues (2003: 834). Here, participants had to monitor the progress of the game and be aware of their past choices and outcomes. This means that they had to use their memory as a platform to construct the context for both the game and each choice within it. Our findings do not support the assumption that the activation of BA10 is significantly stronger for false rather than true claims because of differences in difficulty (Ganis et al., 2003). This is because in our study, both true and false claims could be deceptive, and thus true claims cannot be univocally classified as less difficult.

By contrast, direct comparison of the conditions of claiming falsely and claiming truthfully in this study reveals additional activity mainly in the premotor and parietal cortices (Fig. 4a). These regions are at an even lower level in the executive hierarchy and play a role in selecting and executing specific actions (Hoshi & Tanji, 2000; Schumacher et al., 2007). The greater activity associated in our experiment with claiming falsely is consistent with the need to select the claim for a particular dice throw from the complete range of possible throw configurations. Moreover, given our design and randomized order of dice presentation, the premotor activation is unlikely to result from the difference in the

number of button presses between true and false claims. This is confirmed by the behavioural results (see above) which reveal no differences in response times between true and false claims. Since the response time was recorded after the last button press, this measure is roughly proportional to the number of button presses.

Our results suggest that the truth value of a particular claim is not necessarily a good predictor of the brain activity elicited when people in a specific interactive context decide to be deceptive. The circumstances of the utterance – including e.g. whether the situation is one where deception is an option – are important for understanding the elicited brain activity. Depending on whether one compares brain activity elicited during claiming falsely with claiming truthfully outside a deceptive context (Fig. 4c) or with claiming truthfully within a deceptive context (Fig. 4a), the observed differential brain activity may shift between different levels in a decision-making hierarchy from very 'high' areas that set the context (e.g. BA10) or 'lower' areas involved in concrete action selection among a limited set of very concrete alternatives (e.g. premotor and parietal cortex).

6.2. Responding moves

In Meyer, with regard to the detection of deception, if a player accepts a claim, no direct consequences follow. The game simply continues, although it becomes increasingly difficult to beat previous claims. Challenging the other player's claim is a much more dangerous strategy with an uncertain outcome: the challenger may either win or lose.

When playing games against a human opponent, activity is often observed in a network of regions including medial prefrontal cortex and posterior STS (Gallagher, Jack, Roepstorff, & Frith, 2002; Rilling, Sanfey, Aronson, Nystrom, & Cohen, 2004). Such activity was not observed in the present study. Indeed, when responding moves in the game condition were contrasted with responding moves in the control condition, a decrease in activation in the posterior superior temporal sulcus (pSTS) was evident (see Fig. 5).

This activation is in a very similar location to those observed by Grezes et al. (2004) during a task in which participants attempted to detect whether an actor was pretending that a box they lifted was heavier or lighter than it actually was by observing their movements on video. Grezes observed an increase in activity in pSTS, an area involved in a biological movement, when participants observed actions intended to be deceptive whereas we saw a decrease when participants decided whether or not to accept their opponent's claim. There are a number of critical differences between the two studies. For one thing, the participants in the study of Grezes and colleagues did not directly accuse the actor of deceptive intent. It is probable that activity in pSTS reflects inferential activity concerning a person's intentions from observed movements (Saxe & Kanwisher, 2003; Saxe, Xiao, Kovacs, Perrett, & Kanwisher, 2004). In our study participants could not observe the actions of their opponent although they could see his face on some trials. Our speculation is that the majority of participants did not use observed bodily behaviour to decide whether or not they were being deceived. The likelihood that a player is claiming falsely can also be estimated from the value claimed for the dice throw. Participants may well have suppressed all attempts at metalizing, and focused instead exclusively on their evaluations of probabilities.

These findings suggest that strategies for deception detection in this game are based more on perceived probabilities than on metalizing or moral considerations. Further research is needed to explore whether participants would use metalizing strategies when playing against opponents they know.

6.3. Limitations

This study has attempted to investigate deception in its complexity within a realistic experimental protocol. Of necessity, the ecologically valid paradigm comes at a price. For one thing, we could not control whether and when a subject claimed falsely, or accused the opponent of doing so. Thus, even though roughly balanced, the trials are not quantitatively exactly the same. Together with the relatively small sample size (though it is comparable with other neurological studies of deception), it is possible that activity in some areas of the brain did not survive the tests of statistical significance, and were wrongly rejected as uninvolved. Another critical limitation is that we could not distinguish trials when subjects made true claims with deceptive intent from those in which they did not. In future studies the likelihood of such claims could be modelled, trial-by-trial, on the basis of the dice combinations claimed up to that point of the game.

7. Conclusions

Our study of deceptive behaviour in a popular game suggests that the neural correlates of deception depend very much on context. In this game claiming truthfully can be deliberately deceptive, and activates similar areas to claiming falsely. Our results are in accordance with Koechlin's proposed three level hierarchy of executive control (Koechlin et al., 2003; Koechlin & Summerfield, 2007), and suggest the following relationships: (1) activity in BA10 is associated with setting a deceptive context. (2) The DLPFC/ACC activation that has been found in other neuroimaging investigations of deception is probably indicative of task execution: 'lie' when you are told to. (3) The choice of different false statements is an action-selection process that activates parietal cortex and premotor areas.

A key point emphasized by the Meyer game is that deception typically occurs in the context of an unfolding sequence of events which requires high level strategic planning expressed by a hierarchical executive control. Account must be taken of this in the design and analysis to allow better understanding of the underlying cognitive processes and their neural correlates associated with deceptive behaviour.

Acknowledgements

This work was supported by The Faculty of Humanities at the University of Aarhus, The Danish Research Council for Culture and Communication, and The Danish National Research Foundation's grant to Center of Functionally Integrative Neuroscience.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.neuropsychologia.2010.08.013.

References

- Abe, N., Suzuki, M., Tsukiura, T., Mori, E., Yamaguchi, K., Itoh, M., et al. (2006). Dissociable roles of prefrontal and anterior cingulate cortices in deception. *Cerebral Cortex*, 16, 192–199.
- Abe, N., Suzuki, M., Mori, E., Itoh, M., & Fujii, T. (2007). Deceiving others: Distinct neural responses of the prefrontal cortex and amygdala in simple fabrication and deception with social interactions. *Journal of Cognitive Neuroscience*, 19(2), 287–295.
- Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature reviews*, 7(4), 268–277.
- Baumgartner, T., Fischbacher, U., Feierabend, A., Lutz, K., & Fehr, E. (2009). The neural circuitry of a broken promise. *Neuron*, 64(5), 756–770.
- Braver, T. S., & Bongiolatti, S. R. (2002). The role of frontopolar cortex in subgoal processing during working memory. *Neuroimage*, 15, 523–536.
- Burgess, P. W., Dumontheil, I., & Gilbert, S. J. (2007). The gateway hypothesis of rostral prefrontal cortex (area 10) function. *Trends in Cognitive Science*, 17(7), 290–298.

- 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 of deception: Evidence from activation likelihood estimate meta-analyses. *Cerebral Cortex*, 19(7), 1557–1566.
- Dale, A. M., & Buckner, R. L. (1997). Selective averaging of rapidly presented individual trials using fMRI. *Human Brain Mapping*, 5, 329–340.
- Friston, K. J., Holmes, A. P., Poline, J. B., Grasby, P. J., Williams, S. C., & Frackowiak, R. S. (1995). Analysis of fMRI time-series revisited. *Neuroimage*, 2(1), 45–53.
- Frith, C. D. (2000). The role of dorsolateral prefrontal cortex in the selection of action as revealed by functional imaging. In S. Monsell, & J. Driver (Eds.), *Control of cognitive processes* (pp. 549–565). Cambridge, Mass: MIT Press.
- Gallagher, H. L., Jack, A. I., Roepstorff, A., & Frith, C. D. (2002). Imaging the intentional stance in a competitive game. *Neuroimage*, 16(3 Pt 1), 814–821.
- Ganis, G., 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(8), 830–836.
- Greely, H. T., & Illes, J. (2007). Neuroscience-based lie detection: The urgent need for regulation. *American Journal of Law and Medicine*, 33(2/3), 377–411.
- Greene, J. D., & Paxton, J. M. (2009). Patterns of neural activity associated with honest and dishonest moral decisions. *PNAS*, 106(30), 12506–12511.
- Grezes, J., Frith, C., & Passingham, R. E. (2004). Brain mechanisms for inferring deceit in the actions of others. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 24(24), 5500–5505.
- Henson, R., & Friston, K. J. (2007). Convolution models for fMRI. In Friston, et al. (Eds.), *Statistical parametric mapping* (pp. 178–193). London: Academic Press.
- Hoshi, E., & Tanji, J. (2000). Integration of target and body-part information in the premotor cortex when planning action. *Nature*, 408(6811), 466–470.
- Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture of cognitive control in human prefrontal cortex. *Science*, 302, 1181–1185.
- Koechlin, E., & Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. *Trends in Cognitive Sciences*, 11(6), 229–235.
- Koechlin, E., & Hyafil, A. (2007). Anterior prefrontal function and the limits of human decision-making. *Science*, 318, 594–598.
- Kozel, F. A., Johnson, K. A., Mu, Q., Grenesko, E. L., Laken, S. J., & George, M. S. (2005). Detecting deception using functional magnetic resonance imaging. *Biological Psychiatry*, 58(8), 605–613.
- Langleben, D. D., Schroeder, L., Maldjian, J. A., Gur, R. C., McDonald, S., Ragland, J. D., et al. (2002). Brain activity during simulated deception: An event-related functional magnetic resonance study. *Neuroimage*, 15(3), 727–732.
- Langleben, D. D., Loughhead, J. W., Bilker, W. B., Ruparel, K., Childress, A. R., Busch, S. I., et al. (2005). Telling truth from lie in individual subjects with fast event-related fMRI. *Human Brain Mapping*, 26(4), 262–272.
- Mohamed, F. B., Faro, S. H., Gordon, N. J., Platek, S. M., Ahmad, H., & Williams, J. M. (2006). Brain mapping of deception and truth telling about an ecologically valid Situation: Functional MR imaging and polygraph investigation—Initial experience. *Radiology*, 238(2), 679–688.
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24, 167–202.
- Rammani, N., & Owen, A. M. (2004). Anterior prefrontal cortex: Insights into function from anatomy and neuroimaging. *Nature Reviews Neuroscience*, 5(3), 184–194.
- Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2004). The neural correlates of theory of mind within interpersonal interactions. *Neuroimage*, 22(4), 1694–1703.
- Roepstorff, A., & Frith, C. (2004). What's at the top in the top-down control of action? Script-sharing and 'top-top' control of action in cognitive experiments. *Psychological Research*, 68(2–3), 189–198.
- Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: the role of the temporo-parietal junction in "theory of mind". *Neuroimage*, 19(4), 1835–1842.
- Saxe, R., Xiao, D. K., Kovacs, G., Perrett, D. I., & Kanwisher, N. (2004). A region of right posterior superior temporal sulcus responds to observed intentional actions. *Neuropsychologia*, 42(11), 1435–1446.
- Schumacher, E. H., Cole, M. W., & D'Esposito, M. (2007). Selection and maintenance of stimulus-response rules during preparation and performance of a spatial choice-reaction task. *Brain Research*, 1136(1), 77–87.
- Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge: Cambridge University Press.
- Sip, K. E., Roepstorff, A., McGregor, W., & Frith, C. D. (2008a). Detecting deception: The scope and limits. *Trends in Cognitive Sciences*, 12(2), 48–53.
- Sip, K. E., Roepstorff, A., McGregor, W., & Frith, C. D. (2008b). Response to Haynes: There's more to deception than brain activity. *Trends in Cognitive Sciences*, 12(4), 127–128.
- Sip, K. E. (2009). Neuroimaging of deception in social interaction and lie-detection. *Unpublished Ph.D. thesis*. Aarhus University.
- Spence, S. A., Hunter, M. D., Farrow, T. F., Green, R. D., Leung, D. H., Hughes, C. J., et al. (2004). A cognitive neurobiological account of deception: Evidence from functional neuroimaging. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 359(1451), 1755–1762.
- Spence, S. A., Kaylor-Hughes, C., Farrow, T. F., & Wilkinson, I. D. (2008). Speaking of secrets and lies: The contribution of ventrolateral prefrontal cortex to vocal deception. *Neuroimage*, 40(3), 1411–1418.