Speaking of secrets and lies: The contribution of ventrolateral prefrontal cortex to vocal deception

Sean A. Spence,⁎ Catherine Kaylor-Hughes, Tom F.D. Farrow, and Iain D. Wilkinson

Introduction

Deception has been a perennial feature of human communication and one that remains hard to diagnose reliably (Vrij, 2000; The Global Deception Research Team, 2006). Phenomenologically, deception implicates multiple cognitive processes, especially aspects of prefrontal executive function (Spence, 2004). A liar must attempt to do several things simultaneously: fabricate a plausible narrative (incorporating ‘lies’), selectively conceal the truth, maintain the consistency of their account, control their emotions, monitor the behaviour of their potential victim and, at some level, understand the consequences and potential risks of their deceit (Adenzato and Ardito, 1999). Attempting to lie also suggests that the liar has some understanding of so-called ‘theory of mind’: he realizes (implicitly or explicitly) that different people may hold different beliefs about the same subject and that one party might induce another to adopt a false belief (Gallagher and Frith, 2003). Hence, even a simple ‘white lie’ can comprise a complex concatenation of cognitive, motor and affective processes (Spence, 2004).

Non-human primate, human neurodevelopmental and neuropsychiatric literatures suggest that deceptive capabilities emerged at specific points in the course of primate evolution (Byrne and Corp, 2004), undergo elaboration during individual (human) brain development (e.g., O’Connell, 1997) and, in their most sophisticated manifestations, implicate the prefrontal cortices (Spence et al., 2004). A convergent social psychological literature often invokes the concept of behavioural control: the liar reduces their bodily movements, responds more slowly and speaks less fluently (Vrij, 2000; Vrij and Mann, 2001; Spence et al., 2001; Hughes et al., 2005; Nunez et al., 2005). These effects may be most prominent when the lie is complex, thereby incurring greater cognitive load (Vrij and Mann, 2001).

Current functional neuroimaging methods applied to the investigation of deception rely upon the construction of paradigms that contrast lying with truthful responding (in order to generate a ‘cognitive subtraction’). Hence, a subject might be asked to press a ‘no’ button instead of a ‘yes’ button when answering questions about their recent conduct, denying an item is in their possession or their knowledge of a recently witnessed mock-crime (e.g., Spence et al., 2001; Langleben et al., 2002; Kozel et al., 2005). To date, some authors have been particularly interested in studying subjects telling lies about their pasts (Spence et al., 2001, 2004; Lee et al., 2002, 2005; Ganis et al., 2003; Nunez et al., 2005; Abe et al., 2006, 2007), while others have been more concerned with lies about the present, e.g., scenarios enacted during or immediately prior to

Behavioural and functional anatomical responses exhibited by humans support the hypothesis that deception involves the prefrontal executive. Functional neuroimaging studies have demonstrated that ventrolateral prefrontal cortex (among other areas) is activated during lying, compared with telling the truth. However, despite some consistencies discernible across studies, problems remain concerning experimental validity, e.g., the expediencies of experimenter-sanctioned cue-deception (i.e., subjects being told when to lie); such ‘lies’ may not have comprised adequate proxies for ‘real-life’ deception. In this experiment, we attempted to address such confounding issues by designing an fMRI paradigm in which subjects chose when to lie (thereby minimising cue-dependency), using spoken words, concerning intimate material, which they regarded as ‘embarrassing’; and where further control conditions required them to ‘comply’ with their examiners or to ‘defy’ them (by withholding pre-specified responses). The main effect of lying revealed significant activation of ventrolateral prefrontal cortices. These results replicate and extend our previous findings to those circumstances under which subjects are allowed to choose when to deceive.

© 2008 Elsevier Inc. All rights reserved.

Available online on ScienceDirect (www.sciencedirect.com).

'E Control your mouth...keep to the truth, don't exceed it'

Ptahhotep, 3rd Millennium BC

Corresponding author. Fax: +44 114 22 61522.
E-mail address: s.a.spence@sheffield.ac.uk (S.A. Spence).

Available online on ScienceDirect (www.sciencedirect.com).

doi:10.1016/j.neuroimage.2008.01.035
scanning (Langleben et al., 2002, 2005; Davatzikos et al., 2005; Kozel et al., 2004a,b, 2005; Phan et al., 2005; Mohamed et al., 2006; Gamer et al., 2007). Given the relative novelty of this field, it is impressive that consistencies have actually begun to emerge across studies and laboratories (see Spence et al., 2004; Hughes et al., 2005 for syntheses; and Wolpe et al., 2005; and Spence, in press, for critiques). Most have demonstrated increased activation of prefrontal regions during lying, although the foci implicated have varied between ventrolateral prefrontal (Spence et al., 2001, 2004; Kozel et al., 2004a,b, 2005; Phan et al., 2005; Nunez et al., 2005; Lee et al., 2005; Abe et al., 2006), anterior cingulate (Spence et al., 2001; Langleben et al., 2002; Ganis et al., 2003; Kozel et al., 2004a,b, 2005; Nunez et al., 2005; Lee et al., 2005; Abe et al., 2006) and dorsolateral prefrontal cortices (Lee et al., 2002, 2005; Ganis et al., 2003; Phan et al., 2005; Nunez et al., 2005; Abe et al., 2006, 2007). Most have failed to find brain areas that exhibit greater activation during truth-telling (compared with lying), suggesting that ‘truthfulness’ comprises a relative baseline in human cognition and communication; the truth resembling a pre-potent response, which must be suppressed (Spence, 2004). Our own data implicate the ventrolateral prefrontal cortices (VLPFC) as regions intimately involved in such response suppression, specifically withholding the (pre-potent) truthful response in the context of a lie (see Spence et al., 2001, 2004 and Spence, 2004). While our own evidence derives from experiments where answers have comprised the opposite of the truth (e.g., ‘no’ instead of ‘yes’), Abe and colleagues (2006) found right VLPFC to be activated when subjects lied about what they knew (at a meta-level: ‘I don’t know’ versus ‘I know’), while Kozel’s (2004b, 2005) studies implicate both left and right VLPFC in experiments where subjects falsely denied or affirmed recognition of a salient object. Hence, the truth content withheld by subjects in these experiments has varied greatly between them but what they have shared is the necessity that information be concealed. However, in another context it may be that motor activity is suppressed: it is of interest that left VLPFC activation accompanied intentionally feigned (compared with hypnotically induced) paralysis in one study (Ward et al., 2003), and also occurred during intentionally feigned motor weakness (but not hysterical conversion disorder) in another (Spence et al., 2000). Response suppression may therefore provide a possible means of deceit during both verbal and non-verbal motor behaviours.

However, notwithstanding these different approaches and their convergent findings, functional neuroimaging studies have also shared confounding factors inapplicable to ‘real world’ deception, detracting from their ecological validity. For instance, they are all essentially ‘collaborative’ in nature: the subject consents to be studied and to tell ‘lies’ for their experimenter (Roepstorff and Frith, 2004). Indeed, as well as being sanctioned by the experimenter, lies are usually cued: there is a rule about ‘when’ to lie, so that the subject’s volition is considerably constrained (albeit willingly so, given that they volunteered to participate). Additionally, the subject matter probed during these procedures has often been neutral in content (so far, no one has been asked to divulge or withdraw potentially embarrassing autobiographical material). While Abe and colleagues (2006, 2007) have approached this issue, by asking subjects how they felt about their attempted deceit (did they experience guilt or sinfulness?), the subject of deceit was itself semantic information (i.e., while deceit might have been associated with embarrassment, the material to be concealed was not in itself intrinsically embarrassing).

Obviously, in any new area of research, it is likely that investigators will begin with relatively simple and consensual procedures (indeed, ethical research practice demands that the latter condition always be met) but there is nevertheless a pressing need for greater ecological validity among functional neuroimaging studies of deception.

This study attempted to address some of these limitations (our own studies included): first, by allowing subjects to choose whether to lie or to tell the truth (i.e., lies were not ‘cued’); second, by requiring subjects to answer autobiographical questions concerning events from their pasts which they might genuinely have wished to conceal from others (events which might, therefore, have evoked lies under ‘real-life’ conditions); third, by allowing subjects to respond vocally, with spoken words (rather than button-presses; a technique previously reported in PET studies by Abe and colleagues, above, but not reported in fMRI studies to date (although Ganis et al., 2003, presented a conjunction of combined manual and vocal response data)); and fourth, by the inclusion of further contrast (control) conditions, requiring subjects to choose whether to comply with an examiner’s instructions or to ‘defy’ him by declining to repeat a word (thereby controlling for simple withholding of pre-specified responses). Because of the nature of the scenarios probed during our paradigm we deliberately restricted our subject sample to professional people, all of whom were known to us, and who trusted us with respect to their personal material. [We have previously reported initial findings from a pilot of this method in abbreviated form (relating to 7 subjects; in Spence et al., 2004).]

We hypothesised that lying compared with truth telling would be associated with greater activation of ventrolateral prefrontal (hypothesis 1) and anterior cingulate cortices (hypothesis 2) (Spence et al., 2001, 2004). Furthermore, we hypothesised that no brain areas would exhibit greater activation during truth telling than lying (hypothesis 3) (Spence, 2004).

Method

Demographics and paradigm design

We used an event-related fMRI paradigm to study 17 healthy volunteers (9 females) aged 22–49 years (mean ± SD, 29.9 ± 7.1 years) all of whom were right-handed (Oldfield, 1971). Three further subjects were studied but subsequently excluded because of data quality issues (i.e., intra-scan head movements > 2 mm). All subjects provided written informed consent.

Prior to their study day, each subject prepared written accounts of two true scenarios (A and B) purported to have occurred in their lives. In each case, the subject matter comprised an event that a typical person would probably have wished to conceal from others. We specified that these scenarios should be sufficiently memorable to allow rapid, truthful (and, hence, through elaboration, deceptive) responding inside the scanner. All accounts were written down and placed in sealed envelopes for the duration of the study. The following is a much abridged (and anonymized) example:

The subject had gone to a busy public house in a Scottish city, to meet some friends. After drinking, they had become loud and boisterous. The subject was sitting close to a cigarette machine. He heard someone ask him where the cigarette machine was and he answered sarcastically, implying that it was obvious where it was. Then he turned around to find that the man asking him the question was blind.
Each subject underwent 2 scanning sessions on the same day, one scan relating to each scenario: A and B (hence, there were 34 scans in total; 17 subjects × 2 scenarios each). All trial-stimuli comprised either ‘questions’, probing these target scenarios (e.g. ‘Did you arrive late?’), or balanced ‘instructions’, independent of the scenarios (e.g. ‘Will you now say ‘late’?’). To achieve approximate equivalence between the perceptual features of these stimuli, each stimulus comprised five syllables and the last, ‘operator’, word (e.g. ‘late’ in the former example) always comprised a single syllable. Operator words were matched across question and instruction conditions.

Under the ‘question’ condition, subjects were required to reply either ‘yes’ or ‘no’ in order to respond truthfully or deceptively. Subjects were free to choose when to tell the truth or to lie, but we requested that they should achieve an approximate balance between the total number of truthful and deceptive responses over the course of each scan (without actually counting). Questions were constructed in such a way as to be applicable to any past event; a pilot experiment demonstrated that this was feasible (Spence et al., 2004).

Under the ‘instruction’ condition, subjects were free to ‘comply’ by repeating the ‘operator’ word (i.e., ‘late’ in the above example) or ‘defy’ by saying ‘no’ (our rationale being that both ‘lie’ and ‘defy’ involve withholding a pre-specified response: the ‘correct’ memory or ‘correct’ operator word). Again, we requested approximately equal numbers of ‘comply’ and ‘defy’ responses over the course of each scan.

Hence, there were 2 equal sets of experimental stimuli (36 ‘questions’ and 36 ‘instructions’), administered during each scan (run) and 4 possible experimental responses (2 comprising truth or deception; 2 comprising compliance or defiance). While the experimenters had determined the stimulus frequencies (by matching ‘question’ and ‘instruction’ frequencies, respectively), the relative frequencies of individual response categories depended upon subjects’ performances. Hence, while the frequency of responses related to deception (the sum of ‘truths’ and ‘lies’) equaled those related to defiance (the sum of ‘comply’ and ‘defy’ responses), the final distributions of response type within each category were determined by the subjects’ responses.

MRI parameters

Whole head functional MRI scans were performed using a 1.5-T system (Eclipse, Philips Medical Systems, Cleveland, OH), at the University of Sheffield. Gradient-recalled echo-planar imaging (EPI) was carried out over 72 time points, in which 27×4 mm transverse slices were acquired parallel to the line between the anterior and posterior commissures [echo time (TE)=50 ms, acquisition time (TA)=3 s, repetition time (TR)=6 s, field of view (FOV)=240 mm, in-plane matrix=128×128].

Scanning protocol

Trial-stimuli (questions and instructions) were presented in a single male voice, digitally recorded at 44.1 kHz and 16 bits, and presented binaurally over electrostatic headphones (Resonance Technology Inc., Los Angeles, CA), at an approximately 50-dB sensation level, during interleaved silent periods inside the scanner bore (Fig. 1). Each silent period lasted 3 s. Subjects’ spoken responses were recorded using a microphone attached to the headphone set (subjects were aware that their spoken responses would be listened to by third parties). Scanning sessions comprised 72 consecutive trials, occurring at 6-s intervals. In an event-related design, the series of 36 questions and 36 instructions relating to the four possible responses were counterbalanced throughout each of the 2 experimental runs for each subject.

Following scanning, subjects completed a questionnaire, indicating the truthful answer to each question for both of their scenarios. This was verified against the written scenario retained within their sealed envelopes, allowing the re-coding of each memory ‘question’ answer as either a ‘truth’ or a ‘lie’. Recorded responses to the ‘instructions’ were coded as either ‘comply’ or ‘defy’. These data were used as behavioural regressors in functional image analyses.

It should be noted that although there is no unequivocal ‘ground truth’ in this experimental paradigm (we did not verify the scenarios provided by each subject with a co-informant), we did verify each account in terms of its internal consistency (the written scenario compared with the coded responses discussed at debriefing).

Analysis

We used Statistical Parametric Mapping (SPM2) (www.fil.ion.ucl.ac.uk; Friston et al., 1995), implemented in MATLAB (Mathworks Inc., Sherborn, MA) on a PC, to analyse our data. Data were first realigned to the first image, then normalized to the SPM2 EPI template (2 mm) and smoothed with a 10-mm Gaussian kernel. We did not use a high pass filter (in view of data being obtained over interleaved periods of silence and scanner activity, with a relatively long TR) nor did we use AR1 correction for serial correlations (in view of our intended use of mixed-effects analyses). Three-dimensional coordinates obtained by SPM were interpreted within the Montreal Neurological Institute (MNI) atlas space (using the WFU_Pickatlas tool version 2.0 of Joseph Maldjian MD, Functional MRI Laboratory, Wake Forest University School of Medicine).

An event-related, fixed-effects analysis was used to analyse individual subject data (n=17) at the first-level and a matrix designed to combine both scans (A and B); a random-effects model was used at the second-level. Each subject’s scans were combined and all permutations of conditions convolved with the blood oxygen level-dependent (BOLD) response data. Statistical outcome was calculated using the standard error of the mean on a voxel-wise basis for each area of activation and a parametric map of brain volume in MNI stereotactic space. This method allowed the production of
Results

Behavioural

Analysis of intra-scanner responses revealed that >98% of responses for each of our 17 subjects could be satisfactorily coded.

Table 1 Areas exhibiting relative activation during Lie—Truth and Defy—Comply, and their inverse contrasts

<table>
<thead>
<tr>
<th>Area</th>
<th>Hemisphere</th>
<th>MNI coordinates</th>
<th>Voxels</th>
<th>z</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lie—Truth</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inferior frontal gyrus</td>
<td>L</td>
<td>−34 34 6 75</td>
<td>3.81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle occipital gyrus</td>
<td>L</td>
<td>−30 86 16 38</td>
<td>3.72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cuneus</td>
<td>L</td>
<td>−10 96 16 78</td>
<td>3.97</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cerebellum</td>
<td></td>
<td>−4 −46 −18 79</td>
<td>3.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inferior frontal gyrus</td>
<td>R</td>
<td>32 36 6 57</td>
<td>3.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inferior frontal gyrus</td>
<td>R</td>
<td>60 10 18 46</td>
<td>4.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Putamen</td>
<td>R</td>
<td>20 10 −10 43</td>
<td>3.57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Truth—Lie</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No activations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Defy—Comply</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No activations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comply—Defy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inferior frontal gyrus</td>
<td>L</td>
<td>−26 28 4 5</td>
<td>3.52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inferior frontal gyrus</td>
<td>L</td>
<td>−46 26 0 27</td>
<td>3.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parahippocampal gyrus</td>
<td>L</td>
<td>−20 −14 −22 28</td>
<td>4.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medial frontal gyrus</td>
<td></td>
<td>2 56 6 8</td>
<td>3.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parahippocampal gyrus</td>
<td>R</td>
<td>18 −24 −20 101</td>
<td>3.66</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All are $p<0.001$ uncorrected, 5 voxels extent threshold.

Fig. 2. Mean percentage of responses for each response category over both scan runs (sum of 4 columns=100%, sum of truth and lie=50%; sum of comply and defy=50%).

contrast images for all tasks; these were then used in a second-level analysis to contrast the active and baseline conditions in a between-subjects comparison, i.e., a second level (random-effects) model was created using the contrast images from the first-level analysis. An uncorrected statistical threshold of $p<0.001$, height threshold at 5 voxels extent was applied. It should be noted that our a priori hypotheses related solely to bilateral ventrolateral prefrontal and anterior cingulate cortices (hypotheses 1 and 2), on the basis of our foregoing, published findings (Spence et al., 2001, 2004).

Functional anatomical

There are many potential analyses that might be performed upon a dataset comprising 4 experimental conditions; however, only 2 are directly germane to our primary concern: the functional anatomy of deception. These are:

1. ‘Lie—Truth’, and
2. ‘(Lie—Truth)—(Defy—Comply)’.

Hence, these analyses provide the focus for our report.

First, we performed second-level analyses of the simple contrasts constituting the comparisons: lie minus truth, defy minus comply, and their opposites (Table 1). As hypothesized (hypothesis 1), lying was associated with greater activation of bilateral ventrolateral prefrontal regions, with symmetrical local maxima. These foci were located in Brodmann areas (BA) 45 and 47 (Fig. 3). However, we did not detect significant activation of anterior cingulate cortex during deception (hypothesis 2) (Table 1; although there was a minor focus of less than 5 voxels extent). In keeping with our other a priori hypotheses related solely to bilateral ventrolateral prefrontal and anterior cingulate cortices (hypotheses 1 and 2), on the basis of our foregoing, published findings (Spence et al., 2001, 2004).

Table 2 Areas exhibiting relative activation during lying (compared with truth and defy—comply conditions [(Lie—Truth)−(Defy—Comply)])

<table>
<thead>
<tr>
<th>Area</th>
<th>Hemisphere</th>
<th>MNI coordinates</th>
<th>Voxels</th>
<th>z</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Middle frontal gyrus</td>
<td>L</td>
<td>−28 44 −16 282</td>
<td>3.97</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inferior frontal gyrus</td>
<td>L</td>
<td>−42 30 −4</td>
<td>3.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inferior frontal gyrus</td>
<td>L</td>
<td>−26 28 4</td>
<td>3.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle temporal gyrus</td>
<td>L</td>
<td>−58 −26 −4</td>
<td>4.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cuneus</td>
<td>L</td>
<td>−8 −98 24 31</td>
<td>3.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle occipital gyrus</td>
<td>L</td>
<td>−20 −106 4</td>
<td>4.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thalamus</td>
<td>L</td>
<td>−20 −24 6 56</td>
<td>3.96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle frontal gyrus</td>
<td>R</td>
<td>38 46 −18 5</td>
<td>3.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inferior frontal gyrus</td>
<td>R</td>
<td>62 8 18 14</td>
<td>3.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Superior occipital gyrus</td>
<td>R</td>
<td>38 −86 22 68</td>
<td>3.62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thalamus</td>
<td>R</td>
<td>14 −26 8 16</td>
<td>3.31</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$p<0.001$ uncorrected, 5 voxels extent threshold.

Responses occurred in each response category (mean±SD: 17.95±4.83; where 18 would have comprised a ‘perfect’ response frequency for each of the 4 possible responses, i.e., 25% of the total 72 for each scan run; see Fig. 2). However, there were significantly fewer ‘defy’ responses (one-way ANOVA, $df=3$, $F=35.4$; $p<0.001$; post hoc paired $t$ test, comply versus defy, $df=33$, $t=8.7$; $p<0.001$; Fig. 2).
hypothesis (3, above), we found no areas where truth telling was associated with greater activity than lying (Table 1).

Second-level analyses of the primary contrasts where truth, comply, and defy regressors were subtracted from those associated with the lie response revealed greater activation in bilateral (although especially left) ventrolateral prefrontal cortices (Table 2). It should be noted that the left-sided activation was extensive, comprising over 282 contiguous voxels (local maxima were localised to BA 45/47/11; see Fig. 4 and Table 2).

Discussion

We used fMRI to study the neural correlates of vocal lying in subjects who were allowed to choose when to lie or to tell the truth and when to obey instructions or to defy them. Their lies and truths concerned purportedly embarrassing memories; data which might typically form the basis of lies told in everyday life. In keeping with our first a priori hypothesis (Spence et al., 2001, 2004), we found vocal lying (cf. truthfulness) to be associated with greater activation of symmetrical foci in bilateral ventrolateral prefrontal cortices (VLPFC; regions corresponding to Brodmann areas, BA 45 and 47; Fig. 3). In a subsequent analysis, incorporating ‘defy’ and ‘comply’ conditions, we found such deception-related activation to be most marked in left VLPFC (albeit subsumed within a more extensive activation focus, incorporating several adjacent Brodmann areas). However, this apparently lateralised finding is likely to be confounded through the incorporation of the relatively lateralised ‘comply-related’ activity data (i.e., the ‘comply’ condition was itself associated with solely left-sided VLPFC activity; Table 1). Therefore, the contribution of VLPFC to deception is most cautiously interpreted as being bilateral in distribution.

Study limitations

One of the features of this study was the freedom granted subjects to choose how to respond. While this added ecological validity to our approach, it came at a price: quantitatively uneven (biased) response profiles (Fig. 2). The distribution of responses may have made it more likely that certain of our cognitive subtractions would yield significant results. While this actually served to reinforce the third of our a priori hypotheses [where more numerous ‘defy’ responses might be causally implicated in the withholding of pre-specified responses). We can only speculate whether fewer ‘comply’ and more ‘defy’ responses might have allowed us to identify specific correlates of the latter condition (i.e., we cannot exclude a type II error with respect to the functional anatomy of ‘defiance’). Also, as noted above, the relative excess of ‘comply’ responses likely added to the extent to which the left VLPFC focus of activation predominated on our second analysis (as ‘lie’ and ‘comply’ conditions both activated left VLPFC; see Tables 1 and 2; and Fig. 4).

There are several other weaknesses in the present study. First, while we wished to provide subjects with increased latitude over when to lie or to tell the truth, the requirement that they attempt to balance their response frequencies nevertheless introduced a confounding instruction bias, i.e., they were not totally free to lie or to tell the truth at any frequency they chose. Second, our sample size, though comparable with those of previous studies, is nevertheless relatively small. Third, while our use of a random-effects model increases our confidence regarding the central finding and its applicability to other subjects, drawn from the same ‘population’, the population itself is, of course, highly selected (a sampling bias). We deliberately set out to study a rather homogenous group of professional people, known personally to the experimenters and able to trust them (us) with their personal details (i.e., their scenarios A and B). Hence, while we believe our central finding to be accurate (for this population), we must be cautious before extrapolating it to other social groups.

Fourthly, we must also concede that we have no data regarding ‘ground truth’, the central veracity of our subjects’ embarrassing memories. We did not interview co-informants and did not seek documentary evidence. So we are heavily reliant upon the truthful identification of relevant (and accurate) memories by our subjects. Nevertheless, subjects were required to remain consistent with their scenarios over time and throughout the coding of scores post-scanning, and we also examined the written records of their alleged activities. Indeed, it is striking that the cognitive subtractions involving mnemonic scenarios (‘lie—truth’ and ‘truth—lie’) yielded data so consistent with our a priori hypotheses (1 and 3, above; bilateral VLPFC activation in the former and no areas of greater activation in the latter; Spence et al., 2001).

Nevertheless, in contrast to some previous studies, and our second a priori hypothesis, we did not detect marked activation of anterior cingulate cortex (ACC) in either the simple ‘lie—truth’ or more complex [(lie—truth)−(defy—comply)] contrasts. While there is a small signal in the former comparison (<5 voxels) this was absent from the latter. This suggests either a failure of replication (a problem of reliability) or the possibility that previous reports of ACC activation might have been attributable to certain procedural aspects of those earlier protocols, and not deception per se (a problem of validity). One (congruent) suggestion has been that lies that are ‘spontaneous’ and isolated from any over-arching narrative are more likely to activate ACC, as are lies expressed through manual (cf. verbal) responses (Ganis et al., 2003); all these features would accurately characterize the lies elicited in our first study of deception (Spence et al., 2001) where ACC activation occurred; whereas none applies in the current context (where lies were related to long-remembered scenarios and expressed vocally). Hence, our current emphasis upon vocal lies, describing real events retrieved from subjects’ pasts, might have served to limit ACC’s role in response generation. [Elsewhere, we have provided a detailed critique of the failures of replication seen across earlier fMRI studies of deception, with respect to ACC (and inferior parietal lobule; Spence, in press).]
Nevertheless, given our *a priori* predictions, and an experimental design which necessarily gave subjects a degree of choice over their responses, responses that concerned material of obvious personal salience, we think it striking that VLPFC has once more emerged as prominently activated during the telling of lies (as in Spence et al., 2001, 2004; Kozel et al., 2004a,b, 2005; Phan et al., 2005; Nunez et al., 2005; Lee et al., 2005; Abe et al., 2006).

**The role of VLPFC**

What may we deduce from VLPFC’s activation? As a (very) general rule, the functions of VLPFC can be contrasted with those of dorsolateral prefrontal cortex: the former being especially implicated in the control and suppression of inappropriate cognitions and behaviours (Pandya and Veterian, 1996; Goldberg, 2001; Hooker and Knight, 2006; Aron, 2007), the latter engaged when subjects produce new or complex responses to the environment (Spence and Frith, 1999; Owen et al., 2000). While lesions of lateral orbitofrontal cortex may precipitate dis-inhibition and inappropriate responses among humans, dorsolateral lesions are more characteristically associated with apathy and lack of response initiation; e.g., in transcortical motor aphasia, where the subject may be capable of repeating words spoken to him but incapable of spontaneous speech (Lichteim, 1885).

VLPFC BA 45 is located in the pars triangularis in the inferior frontal gyrus, extending rostrally into the inferior prefrontal convexity (Pandya and Veterian, 1996). It receives projections from virtually all areas of prefrontal cortex (with the exception of DLPC, BA 9/46), together with superior (auditory) and inferior (visual) temporal regions, caudal insular, cingulate and rostral parahippocampal gyri (Pandya and Veterian, 1996). VLPFC (including BA 47) is preferentially activated during response suppression, as illustrated in lesion studies and ‘go—no-go’ tasks in health and disease, and has been implicated in conditional learning (e.g., Aron et al., 2004). Hence, it is involved in the control of behaviour, with especial reference to the withholding of inappropriate responses and learning of new contingencies. A review by Hooker and Knight (2006) emphasizes the contribution of VLPFC to helping to maintain goal-directed behaviours by filtering-out potentially distracting emotional stimuli and perceptions: the emphasis is again upon the inhibitory control of cognition. Activation of VLPFC during our vocal lying paradigm appears congruent with the procedural requirement that our liars suppress the truth. Furthermore, the design of our study allowed us to partially discount simple response reversal (saying ‘no’ for ‘yes’), and merely following rules (confounders that were present in our earlier work; Spence et al., 2001; and which are of relevance, given the contribution of VLPFC to rule-following behaviour *per se*; see Bunge, 2004). Hence, we have greater confidence in VLPFC’s activation being specifically related to deception rather than to other, superficial features of the experimental design.

**Congruence with pathological conditions**

We have previously hypothesized that VLPFC is implicated in deceptive responding since it is crucial that the subject refrain from emitting relatively pre-potent, pre-specified (truthful) responses (Spence et al., 2001). Indeed, several brain imaging studies have revealed VLPFC activation during lying (above). However, is there any evidence that damage to this region impacts the veracity of speech: thereby increasing its ‘truthfulness’? In one of the classic descriptions of ‘pseudo-psychopathic syndrome’, consequent upon orbitofrontal lesions, Blumer and Benson (1975) described patients who became ‘outspoken’, ‘brash and disrespectful’, ‘open and frank,’ and ‘lack(ing) of adult tact and restraints’ (pp. 154–158). These attributes suggest that a patient sustaining such a lesion is rendered pathologically ‘honest’. More recent, experimental, studies have demonstrated that such patients are abnormally prone to sharing intimacies with others, going further than is strictly necessary in the social context (Beer et al., 2003). However, it should be emphasised that these findings relate to orbitofrontal cortex as a whole, albeit with some emphasis upon lateral orbitofrontal or ventrolateral prefrontal cortices, particularly BA 47; they do not seem to favour the left frontal lobe or BA 45.

Another strand of evidence concerns pathological lying. Two recent studies from Raine’s group have demonstrated that antisocial subjects with prominent histories of deception (e.g., fraud) exhibit increased prefrontal white matter relative to appropriate comparator groups, an increase most marked in orbitofrontal cortices (Yang et al., 2005, 2007). Hence, while lesions to orbitofrontal cortex may render patients pathologically truthful (above), an increase of white matter volume in these regions is associated with more frequent deceptive behaviour. Yang and colleagues (2007) posit that their antisocial subjects may share a facility for deceiving others (see Spence, 2005, for alternative interpretations).

**What might lying and compliance share in common?**

A parsimonious theory of deception would suggest that VLPFC supports deceptive behaviour through its role in suppressing truthful responding (Spence, 2004). The present study confirms that its involvement can be demonstrated even when the subject matter probed is potentially embarrassing. It also suggests that VLPFC may be implicated in both deception and social compliance. Though highly speculative, one interpretation might be that experimental lying and compliance are similar in that the subject is ‘saying what the experimenter wants’ (Roepstorff and Frith, 2004). Alternatively, and more specifically, they are both relatively ‘inauthentic’ modes of responding: compared with truthful responding and refusing to repeat the experimenter’s words, lying and compliance both appear less ‘open and frank’ in Blumer and Benson’s (1975) memorable phrase.

Finally, we must acknowledge that the ‘secrets’ that our subjects shared (and withheld), though more intimate than those described in previous studies, are mild by comparison with those that might be probed in forensic settings (e.g., Vrij and Mann, 2001). It seems likely that accurate identification of truly immoral or antisocial acts will require both painstaking empirical progress and thoughtful ethical consideration. We have recently taken tentative steps in this direction (Spence et al., in press).

**Conclusions**

To our knowledge, ours is the first fMRI study to probe subjects’ responses as they tell lies about purportedly embarrassing events from their pasts. We also allowed subjects to ‘choose’ when to lie or tell the truth, thereby reducing the constraints of cued behaviour. While the conduct of our study remained essentially collaborative (subjects agreed to participate, under conditions of anonymity), the selection of subjects who were prepared to trust the experimenters with their personal scenarios allowed us to study purportedly sensitive material (as might be envisaged in putative forensic ‘lie detection’). Our results add to evidence that VLPFC is implicated in the telling of lies; including those that concern intimate material.
Acknowledgments

We thank the subjects who took part in this study, our colleagues in SCANLab and Academic Radiology, and the following for their contribution to pilot studies: Dr Mike Hunter, Dr Venkat Ganesan, Dr David Leung and Dr Russell Green.

References


Adenzato, M., Ardito, R.B., 1999. The role of theory of mind and deontic


Blumer, D., Benson, D.F., 1975. Personality changes with frontal and temporal


Byrne, R.W., Corp, N., 2004. Neocortex size predicts deception rate in


