Neural Correlates of Different Types of Deception: An fMRI Investigation

G. Ganis[1–3], S.M. Kosslyn[1,4], S. Stose[5], W.L. Thompson[1] and D.A. Yurgelun-Todd[6]

[1]Department of Psychology, Harvard University, Cambridge, MA 02138, [2]Department of Radiology, Massachusetts General Hospital Martinos Center, Charlestown, MA, 02129, [3]Harvard Medical School, Boston, MA 02115, [4]Department of Neurology, Massachusetts General Hospital, Boston, MA 02142, [5]Department of Psychology, University of Virginia, Charlottesville, VA 22904 and [6]Brain Imaging Center, McLean Hospital, Belmont, MA 02478, USA

Deception is a complex cognitive activity, and different types of lies could arise from different neural systems. We investigated this possibility by first classifying lies according to two dimensions, whether they fit into a coherent story and whether they were previously memorized. fMRI revealed that well-rehearsed lies that fit into a coherent story elicit more activation in right anterior frontal cortices than spontaneous lies that do not fit into a story, whereas the opposite pattern occurs in the anterior cingulate and in posterior visual cortex. Furthermore, both types of lies elicited more activation than telling the truth in anterior prefrontal cortices (bilaterally), the parahippocampal gyrus (bilaterally), the right precuneus, and the left cerebellum. At least in part, distinct neural networks support different types of deception.

Introduction

Deception occurs when one person attempts to convince another to accept as correct what the prevaricator believes is incorrect [typically in order to gain a benefit or avoid punishment; cf. (Spence *et al.*, 2001)]. Given the obvious importance of detecting deception, individuals as well as entire societies have long sought reliable methods for determining when a person is lying (Ekman, 1992, 2001). Traditionally, observers have tried to detect lies by noting subtle behavioral cues. Indeed, researchers have characterized a number of nonverbal cues that are associated with deception, but none of these cues is entirely diagnostic or reliable. For example, *microexpressions*, brief and incomplete changes in expression – such as head shakings or negative facial expressions – are among the most reliable nonverbal cues for deception (Mehrabian, 1971; Burgoon and Buller, 1994; Frank and Ekman, 1997); the pitch of the voice tends to be elevated when people are engaged in deception (Vrij, 1994; Zuckerman *et al.*, 1979); the body posture is generally more rigid when one is lying than when one is telling the truth (Mehrabian, 1971; Vrij, 1994); and alterations in patterns of eye contact are also associated with deception (Horvath *et al.*, 1994). These behavioral cues are generally thought to reflect increased physiological arousal during deception, which may arise because the individual feels guilty, is afraid of being detected, or is excited at the thought of deceiving others (Ekman, 1992).

In an effort to develop reliable objective measures of deception, researchers and criminologists have devised various machine-based techniques that typically attempt to measure arousal. The *polygraph*, which monitors physiological functions such as heart rate, breathing rate, and skin conductance, has been used in diverse ways in the service of detecting deception (Office of Technology Assessment, 1983). One class of methods (control question test, CQT) compares physiological responses to *relevant* questions (e.g. 'Did you steal the car that was reported missing?') with responses to *irrelevant* questions (e.g. 'Were you born in California?') and *control* questions [e.g. 'During the first 15 years of your life, did you ever steal anything?' (Horowitz *et al.*, 1997; Podlesny and Raskin, 1977)]. Another class of polygraph methods, the Guilty Knowledge Test (GKT), has focused on detecting physiological changes in response to questions that could only be answered by the perpetrator of a crime (Lykken, 1974; MacLaren, 2001).

One general problem with all polygraph methods is that they detect increases in measures that reflect increased arousal, which are typically interpreted as reflecting guilt and fear. These measures can confound lie detection in two ways. First, guilt and fear can occur in many situations other than during deception, and hence the measures do not necessarily index deception *per se*. Second, if the liar does not feel guilty, he or she may not evince the physiological reaction. Thus, even though exceptions have been noted (Raskin and Hare, 1978), in general standard lie detection techniques might be characterized as 'guilt detection techniques'.

In this study we focus on one potential method for circumventing these problems: Namely, we examine directly the organ that produces lies, the brain (other methods, not explored here, may include using behavioral or peripheral psychophysiological measures that correlate with cognitive processes of interest). This neurobiologically based strategy relies on identifying specific patterns of neural activation that underlie deception. The logic we adopt here has led researchers to use brain-monitoring techniques to try to develop improved lie detection techniques. For instance, researchers have used event-related potentials in the Guilty Knowledge Test, and have reported more accurate discrimination rates than is possible with polygraphic methods (Allen and Iacono, 1997). However, even these methods are not ideal because of the limited spatial resolution of the technique, making it difficult to disentangle complex cognitive processes occurring simultaneously. In the present study we used functional magnetic resonance imaging (fMRI) to monitor neural activation while people lied or told the truth. To date, three fMRI studies of deception or a related topic (e.g. malingering) have been published (Spence *et al.*, 2001; Langleben *et al.*, 2002; Lee *et al.*, 2002). The results have not been consistent. Because the precise questions asked and methods employed in these studies are different from those used in our study, we will defer discussing them until after describing our own findings.

In addition to the limitations inherent in previous lie detection techniques, a second – and in some respects deeper – problem with all prior lie detection methods is that they rest on the assumption that there is only one type of lie. One of the strengths of contemporary theory in cognitive neuroscience is that it distinguishes subtypes of a given function. For example, 'memory' may be decomposed into working, episodic, semantic memory, and so on. In the present work, we applied the same approach to study deception and asked what types of processing

Cerebral Cortex Aug 2003;13:830–836; 1047–3211/03/$4.00

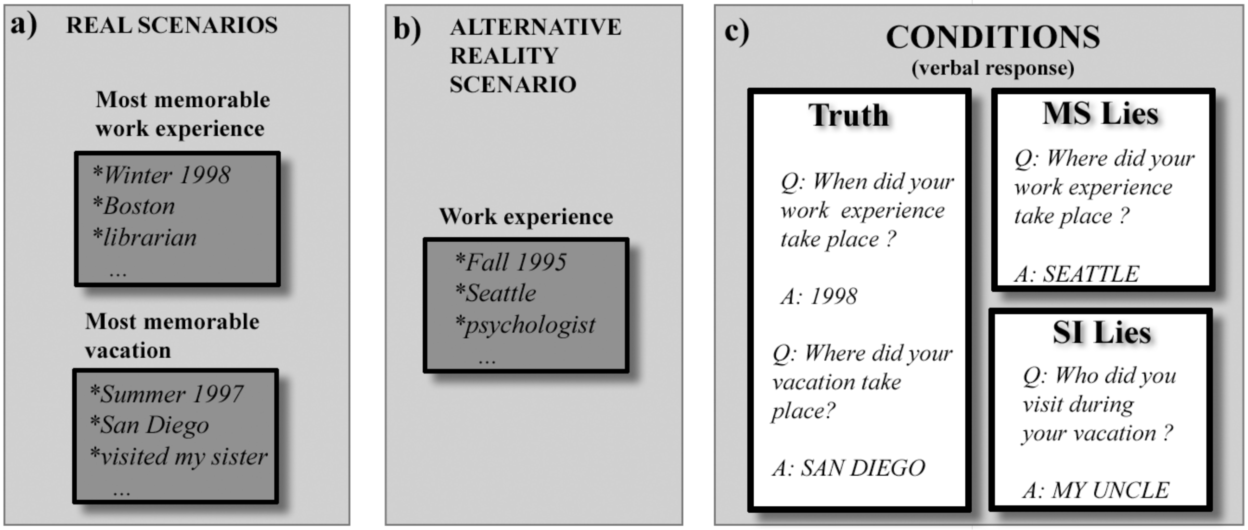

Figure 1. Summary of paradigm. (*a*) During the initial interview participants provided details about their most memorable vacation and work experience; (*b*) during an intensive session preceding fMRI scanning, each participant devised and memorized a coherent alternative scenario for one of the situations ('work experience' in this example); (*c*) during the fMRI session people were asked in separate sets of trials to tell the truth (T), lie according to the memorized alternative scenario (MS), or answer with a spontaneous lie (SI). Note that each condition was also tested in a forced-choice paradigm (not shown here) that required pressing one of two buttons.

differences might distinguish different types of lies. We focused on two orthogonal dimensions. Along the first dimension we differentiated between spontaneous and memorized lies. A spontaneous lie is constructed based on stored information, probably using a mixture of semantic and episodic knowledge. For instance, one could lie about what one ate for lunch by retrieving a specific episode about what one ate some other time (episodic memory), or one could think about what foods one plausibly could eat for lunch (semantic memory).

In contrast, if a lie is memorized in advance, one needs only to retrieve it from memory. We hypothesize that one defining feature of memorized lies is that they are not as rich in detail or as well learned as truthful knowledge. Real experiences are rich in incidental perceptual detail, whereas lies may often consist only of bare-bones descriptions. Moreover, actual experience may be registered in multiple modalities, and hence is subject to 'dual encoding' (Paivio, 1971). If so, then all else being equal, representations of lies should be more difficult to retrieve than truthful knowledge.

In the second dimension, we differentiated between lies that may be isolated versus lies that fit into a scenario (i.e. a coherent story). Spontaneous lies that are isolated are easier to generate than coherent lies because one does not have to cross-check details to ensure that they fit into a larger scheme. In terms of underlying neurocognitive processes, this translates into working memory's being more engaged when one generates a coherent lie than an isolated lie because more information has to be held in mind and evaluated (Smith and Jonides, 1998). In contrast, for memorized lies, those that fit into a coherent scenario may be easier to generate because it is easier to recall a lie when more retrieval cues are present (Schacter, 1996). Treating these two dimensions as orthogonal, we can identify four distinct types of lies — and each type should be associated with a different pattern of neural processing.

In this study we focused on two extremes from this taxonomy: Spontaneous-Isolated (SI) lies and Memorized-Scenario (MS) lies (Fig. 1). To construct a SI lie one needs to retrieve information from semantic and/or episodic memory and generate a viable lie rapidly, keeping in mind many possibilities (i.e. the truth, so to be able to avoid it, together with a number of potential lies) and selecting among them. Accordingly, we hypothesized that, relative to telling the truth, telling SI lies should result in stronger activation in neural structures underlying: (i) semantic and episodic retrieval [e.g. ventrolateral prefrontal cortex and anterior prefrontal cortex, respectively (Duncan and Owen, 2000; Fletcher and Henson, 2001), precuneus (Krause *et al.*, 1999), and possibly ventral stream regions, if these retrieval operations are also accompanied by visual imagery (Kosslyn *et al.*, 2001)]; (ii) working memory [e.g. dorsolateral prefrontal cortex (Smith and Jonides, 1998)]; and (iii) response inhibition and conflict monitoring [e.g. the anterior cingulate (Braver *et al.*, 2001; Ruff *et al.*, 2001)]. In contrast, to generate MS lies, one needs only to access knowledge stored in episodic memory. Thus, for MS lies we hypothesized that we would find increased activation (relative to telling the truth) in brain regions associated with retrieving information from episodic memory [anterior prefrontal cortex (Buckner *et al.*, 1998; Duncan and Owen, 2000) and the precuneus (Krause *et al.*, 1999)].

Materials and Methods

Participants

Three males and seven females, between the ages of 20 and 30 (mean age 25), volunteered to participate for pay. All participants had normal or corrected-to-normal vision. The study was conducted with the informed consent of each participant and the approval of the Harvard University and McLean Hospital Institutional Review Boards.

Pre-testing Procedure and Stimuli

We began by interviewing all participants to obtain detailed information about two of their actual experiences (Fig. 1*a*). We asked them to write down details about the most memorable work experience they ever had ('work' situation) and about the most memorable vacation they ever took ('vacation' situation). Participants returned to the lab on average 1 week after the initial session, which allowed us time to prepare and record questions based on their specific experiences. We first reviewed in general terms the reports they provided earlier, to refresh their memories. Then we instructed the participants to generate an alternative scenario for one of the situations (Fig. 1*b*). The investigator helped them to constrain the scenario so that it fit the prepared questions. For example, if they actually took their vacation in Florida, we instructed them to pretend that the vacation took place in another location in the United States (e.g. California); if they traveled there by car, we instructed them to pretend they used another means of transportation (e.g. by plane), and so on. We helped the participants in this process by ensuring that the scenarios they generated were coherent and internally consistent. Participants were then asked to rehearse and memorize this 'false scenario', so that they could answer questions based on it. This scenario was then used to assess MS lies.

Following this, the participants took part in a practice session outside the scanner, which simulated what they would experience during the actual test session. During this practice session, participants first were given instructions and asked to paraphrase them to ensure that they understood the task. The instructions reiterated the importance of remaining still during each scan; to minimize the possibility of random responses during the SI one-word scans, when participants produced verbal responses, we told them that their responses (without the condition label) would be recorded and reviewed later by a separate 'judge' who would try to determine which were lies. This instruction was not only intended to make them really try to lie, but also to discourage bizarre responses (such as 'purple' or 'heavy' to a question like 'what color are your mother's eyes?', which would be easily spotted as a lie). We then asked whether they had any questions, and if so the investigator answered them.

fMRI Testing Procedure

Following the practice session, the participant was positioned in the magnet and we administered three conditions while using fMRI to monitor brain activation (Fig. 1*c*).

For the MS lie condition, we asked the participants to generate a lie based on the false scenario (e.g. work) that they previously had memorized. For the SI lie condition, we asked them to lie about the other scenario (e.g. vacation), which required them to generate lies in real-time without trying to devise a coherent story, that is without making an effort to generate lies that were consistent with each other. As noted earlier, the instructions emphasized that each lie had to be plausible within the context of the question (i.e. such that an external observer without additional knowledge could not distinguish the lie from a truthful response). Finally, for the Truth condition we asked them to provide honest answers based on the information they provided initially. All the questions used in the MS and SI conditions were new. However, half of the questions in the Truth condition were repeated from the MS and SI conditions. This was a compromise between using different questions for the lie and truth conditions (and running into potential stimulus set confounds) and using the same questions (but running into potential repetition effects). Imaging was conducted in a dark room, and the participants wore headphones, which both helped to attenuate the sound of the scanner and also allowed us to present probes.

Each block began with an auditory cue (e.g. 'Vacation, Truth'), which indicated the condition that would follow. We explained the nature of the cues in advance. In separate blocks, each question required either a binary button-press response ('yes'/'no', to a question such as 'Did you go to Florida on your last vacation?') or a verbal, single-word response (for example, in response to a question such as 'Where did you go on your last vacation?', one could say 'Morocco'). The button-press responses were automatically recorded by the computer, whereas the verbal responses were recorded manually by the experimenter. The participants, however, were told that their verbal responses were tape recorded (and thus they should lie convincingly, to deceive later judges). Thus, we acquired two types of responses for each of three conditions (two types of lies versus truth). To obtain a stable baseline we administered the truth condition twice (with different questions) which resulted in a total of eight blocks (two for the MS lies condition, two for the SI lies condition, and four for the Truth condition; half the blocks requiring a manual response, and half requiring a verbal response for each condition).

Each block lasted 210 s, beginning with 30 s of rest followed by three cycles during which 30 s of an experimental condition (five trials, lasting 6 s each; the auditory question lasted 3 s, on average, and the length did not differ between conditions) alternated with 30 s of rest. Therefore, there were 15 trials for each block, regardless of the response modality. We used a blocked design instead of an event-related paradigm because the study was designed to detect differences between conditions (given the relatively limited number of stimuli), not to estimate the timecourse of the hemodynamic response (Birn *et al*., 2002). We counterbalanced the order of the conditions across participants with the constraint that both lie conditions always came first to avoid potential short-term interference from the truth condition (such as actively having to inhibit primed responses).

Image Acquisition

MRI acquisition was conducted on a 1.5 T scanner (General Electric Signa, Milwaukee, WI) with a standard quadrature head coil and echoplanar capability (Instascan, ANMR Systems, Wilmington, MA). T_2*-weighted echoplanar images sensitive to blood oxygen level- dependent contrast (BOLD) were acquired during the functional scans (gradient echo; T_R = 3000 ms; T_E = 40 ms; alpha = 75°; image matrix = 64 × 128; in-plane resolution = 3.125 × 3.125 mm; slice thickness = 6 mm). Sixteen to twenty axial slices per volume were acquired, depending on head size. Anatomical images for these slices were obtained with a T_1-weighted sequence (T_R = 500 ms; T_E = 11 ms). Whole-brain anatomical images (coronal) were acquired after the functional scans with a SPGR sequence (T_R = 35 ms; T_E = 5 ms; FOV = 240 mm; slice thickness = 1.5 mm; imaging matrix = 256 × 192).

Image Analysis

Data were analyzed with AFNI (Cox, 1996). The data were first corrected for motion artifacts using AFNI program '3dvolreg' (Cox and Jesmanowicz, 1999). Because this motion correction algorithm can only correct small motions, the threshold for the exclusion of a scan due to motion was a shift of more than 4 mm in any direction and a rotation of more than 1.5°. No scan exceeded this threshold. We estimated maps of percent BOLD signal change for the eight series of trials by using the correlation methods described in Cox (Cox, 1996). We then transformed these maps into Talaraich space (Talaraich and Tournoux, 1988), using the scheme provided by AFNI. Briefly, the brain is divided into 12 regions by means of user-placed markers (including the anterior and posterior commissures) and a continuous piecewise affine transformation is then used to transform the original brain into Talaraich space. These maps were then resampled onto a 3 × 3 × 3 mm grid and smoothed with a Gaussian filter (full-width half-maximum = 7 mm, AFNI program '3dmerge').

Following this, we submitted the maps to a one-way repeated-measures ANOVA to identify regions of interest (ROIs) that showed a main effect of condition. We only retained clusters of 40 or more contiguous voxels that were significant at $P < 0.005$, leading to an alpha of 0.05 for the entire 3D image. This minimum cluster size was determined using the Monte-Carlo approach described by Xiong and colleagues (Xiong *et al*., 1995) and implemented by programs '3dFWHM' and 'AlphaSim' (with 1000 iterations) in AFNI. This method (i) estimates the smoothing present in the data based on a variant of the algorithm described by Forman *et al*. (Forman *et al*., 1995), and (ii) determines the number of clusters of a given size that would be significant at a particular threshold due to chance. The probability of a false positive detection across the entire image is then determined by the frequency counts of cluster sizes. Next, we performed planned contrasts on functionally defined ROIs, comparing the two lie conditions with the truth condition, and the two types of lies directly. The alpha for the planned contrasts within the ROIs was $P = 0.005$, corresponding to $Z = 2.81$. Preliminary analyses did not reveal reliable differences in the results from the two response modalities (yes/no and one-word response), and thus we combined the data over the two types of responses in order to increase statistical power.

Results and Discussion

The behavioral data from seven participants (data from the remaining three were not recorded due to equipment problems) indicate that the participants did in fact follow the instructions. In fact, the average error rates (defined as responses that were not appropriate for that condition) were less than 10% in every condition (51 of the 56 participant-by-condition cells had error

Table 1

Twelve ROIs were identified functionally as those regions that exhibited a main effect of condition ($P < 0.05$, corrected at voxel and cluster level). Planned contrasts identified regions that exhibited stronger activation in the Spontaneous-Isolated (SI) and Memorized-Scenario (MS) lie conditions relative to the Truth condition (T), in the SI lie condition relative to the MS condition, and in the MS condition relative to the SI condition

Brain region (Brodmann area)	Coordinates (x, y, z)	Cluster size (μl)	Z value	Contrast			
				SI > T	MS > T	SI > MS	MS > SI
Right middle frontal gyrus (superior BA 10)	31, 51, 24	7182	3.67	*	*		
Right middle frontal gyrus (inferior BA 10)	30, 53, 4	1080	2.85		*		*
Left middle frontal gyrus (superior BA 10)	–36, 46, 21	3024	3.17	*	*		
Anterior cingulate (BA 32)	4, 6, 39	7641	3.13	*		*	
Right middle frontal gyrus (BA 8/9)	26, 25, 38	3510	3.2	*			
Left precentral gyrus (BA 4)	–39, –12, 40	2511	3.39	*		*	
Right pre/postcentral gyrus (BA 3/4)	63, –17, 28	1971	3.25	*		*	
Right fusiform/parahippocampal gyrus (BA 36/37)	23, –46, –10	2025	3.05	*	*		
Left fusiform/parahippocampal gyrus (BA 36/37)	–29, –39, –10	1620	3.49	*	*		
Right cuneus (BA 17/31)	9, –66, 12	1431	2.97	*		*	
Right precuneus (BA 7)	12, –46, 46	2376	3.23	*	*		
Left cerebellum	–4, –71, –10	4808	3.51	*	*		

Coordinates and Z values are for the maximum within each ROI. Volumes were computed after transformation to the Talairach space. The asterisks indicate significant contrasts within the corresponding ROI ($P < 0.005$, uncorrected).

rates lower than 7%, and only two had error rates over 20%, but were still well above chance). The error rates for the different conditions were comparable [$F(5,30) = 1.85$, $P > 0.1$, using the Greenhouse–Geisser correction for inhomogeneity of variance]. Response times (recorded only for the yes/no response conditions) initially appeared slower for the SI (859 ms) and MS (838 ms) conditions than for the T condition (613 ms), but there was no significant difference between conditions [$F(2,12) = 2.54$, $P > 0.1$].

The ANOVA on the BOLD data revealed 12 regions that exhibited a main effect of condition, listed in Table 1. Follow-up analyses revealed that a number of these regions were activated by *both* lie conditions compared to the truth condition (Table 1 and Fig. 2): anterior portions of the middle frontal gyrus bilaterally (superior BA 10, section 4 in Fig. 2), the fusiform/parahippocampal gyrus bilaterally (BA 36, 37), the right precuneus (BA 7), and the left cerebellum.

We next compared the two lie conditions directly. When participants produced SI lies, a number of brain regions were activated more strongly than when they produced MS lies (Table 1 and Fig. 2): the anterior cingulate, extending into the left premotor cortex (BA 32, 6), the left precentral gyrus (BA 4), the right precentral/postcentral gyrus (BA 3/4), and the right cuneus (BA 17).

In contrast, when participants produced MS lies, only the right anterior middle frontal gyrus (inferior BA 10, section 2 in Fig. 2, ~20 mm below the right anterior middle frontal gyrus focus activated in common in both lie conditions) was activated more strongly than during the SI condition.

These results show that different patterns of brain activation arise when people tell lies than when they tell the truth, and the type of lie modulates these patterns. These findings support the idea that lying and telling the truth rely on systematically different neural processes, at least in this paradigm. Furthermore, they suggest that 'lying' is not a single process or function, but instead is a heterogeneous category; therefore, studies aimed at detecting deception may need to examine different neural signatures (or combinations of signatures) to detect the different types of lies. It is important to emphasize that we do not claim that any of these brain regions are specialized for lying; rather, as we have described in the introduction, our neurocognitive framework assumes that the generation of various types of lies engages different combinations of general-purpose cognitive processes which, as an ensemble, may provide reliable neural signatures for various types of lies. These results suggest that during the generation of a SI lie, one may need to access semantic and episodic knowledge. These functional demands were reflected by activation of (bilateral superior) BA 10 (Grady, 1999), the precuneus (Krause *et al.*, 1999) and the cerebellum (Andreasen *et al.*, 1999). In addition, this process appears to be accompanied by visual imagery [right cuneus (Kosslyn *et al.*, 2001)]. These results also suggest that, while constructing a viable lie, the retrieved information is maintained in working memory, which was reflected by activation in BA 8/9 and posterior visual cortex [fusiform gyrus and cuneus (Grady, 1999)]. Moreover, they suggest that, while formulating the lie one may need to check that it is not the truth but nevertheless is plausible, which may in part be responsible for activation in the anterior cingulate (Ruff *et al.*, 2001). Finally, they suggest that the generated lie is encoded into episodic memory, which would explain why the parahippocampal cortex (Wagner *et al.*, 1998; Epstein *et al.*, 1999) was activated.

In contrast, to generate a MS lie, the participants first would have needed to retrieve the false scenario from episodic memory, which relied on bilateral superior BA 10 and right inferior BA 10 (Grady, 1999), the precuneus (Krause *et al.*, 1999), and the cerebellum (Andreasen *et al.*, 1999). They then may have generated a lie according to the memorized scenario, and subsequently encoded the reconstructed lie into episodic memory [which again relied on parahippocampal cortex (Epstein *et al.*, 1999; Wagner *et al.*, 1998)]. Note that, in general, activation of right BA 10 tends to be larger for items that are repeated (old) than for new items (Wagner *et al.*, 1998). Since half of the T questions were repeated from the SI and MS blocks, this could introduce a confound. However, if anything, item repetition should *decrease* the size of the difference between the MS and T conditions. Since none of the questions in the lie conditions were repeated, this factor cannot affect the direct comparison between the MS and SI conditions.

The only region that was significantly more active when participants told MS lies than when they told SI lies was the right inferior BA 10, which has been implicated in episodic retrieval operations [(Fletcher and Henson, 2001; Grady, 1999); see (MacLeod *et al.*, 1998) for other tasks that activate BA 10)]. Although MS lies were both coherent and memorized, the task did not require subjects to cross-check or make use of the

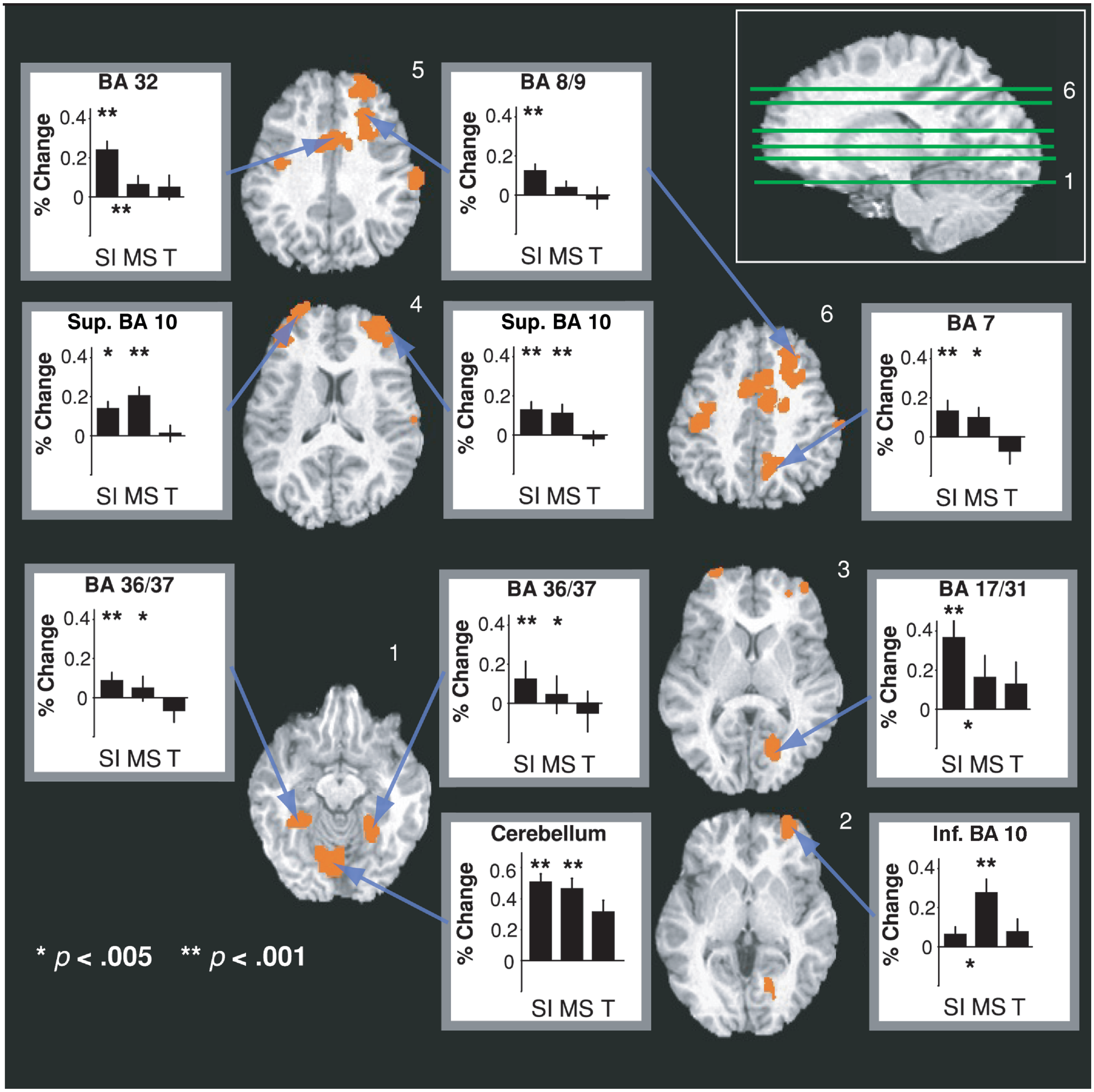

Figure 2. Summary of results. Brain regions that showed a main effect of condition are indicated in orange ($P < 0.005$, cluster size > 40 voxels, $n = 10$). The right side of each axial section corresponds to the right side of the brain. Each bar graph shows the mean and standard error of the activation in the corresponding ROI for the SI, MS and T conditions. The asterisks above the SI and MS bars indicate the respective P values (bottom left) relative to the T condition. The asterisks between the SI and MS bars, below the graph, indicate the P value of the comparison of SI with MS. All Ps are for the maxima within each ROI. For brevity, no graphs are provided for the precentral and postcentral gyrus activations (visible in sections 5 and 6) reported in Table 1.

coherent characteristics of the memorized scenario. Thus, we interpret the findings in terms of the fact that MS lies were based on memorized scenarios *per se*. Specifically, the right inferior BA 10 may have been more active during MS lying because SI lies do not require one to retrieve information only from episodic memory; instead, SI lies may rely on a mixture of episodic and semantic information. The fact that the MS lies were learned as a scenario, which included relations among the lies, may have also been a factor contributing to this effect in BA 10 (Christoff *et al.*, 2001).

In addition, the fact that BA 10 was more strongly activated during MS lying than during telling the truth could indicate that MS lies are more difficult to retrieve than truthful knowledge. That is, unlike MS lies, truthful knowledge is acquired via extensive and multimodal interactions with the real world. Thus, truthful memories are more redundant and have many more retrieval cues than MS lies, which are acquired during a brief verbal exchange that provided only limited details to be encoded. One neural correlate of this difference may be that truthful memories are encoded in a larger network of areas than rehearsed lies, as suggested by modality-specific areas activated during episodic retrieval (Wheeler *et al*., 2000).

A set of regions was significantly more active during the SI than the MS lie condition. The stronger anterior cingulate activation is consistent with the notion that this region is involved in conflict monitoring (Ruff *et al*., 2001) and in the inhibition of competing responses (Braver *et al*., 2001). Although competing responses may also be present when people tell MS lies because of incidental recall of truthful knowledge, the response in the MS lie condition is unique and is entirely determined by the alternative scenario memorized prior to the scanning session. Anterior cingulate activation has also been associated, among other factors, with working memory load (Bunge *et al*., 2001), and with arousal (Lane *et al*., 1997), which could also contribute to the present finding of stronger anterior cingulate activation in the SI condition. Stronger activation in visual cortex is consistent with the idea that visual imagery may be used to generate SI lies. Visual imagery may not be used when one tells MC lies because the participants memorized verbal responses when constructing the alternative scenario. Kosslyn and Jolicoeur (Kosslyn and Jolicoeur, 1980) found that imagery typically is not used spontaneously when people have either memorized the response or can infer it easily from associated information (such as superordinate categories); the SI condition had neither of these characteristics, whereas the MS condition did.

Lastly, we note that some of the activated regions were entirely unexpected. Specifically, we did not predict modulation of activation in the primary motor cortex (close to the hand and mouth representations). We can speculate that these activations may be due to some differences in the motor response (the response times were not different, but one could press the button harder while generating an SI lie), or could be related to the presence of competing potential lies, because they were largest in the SI condition (DeSoto *et al*., 2001).

To our knowledge, there have been only three published fMRI studies related to deception (Spence *et al*., 2001; Langleben *et al*., 2002; Lee *et al*., 2002). In the study by Spence *et al*. (Spence *et al*., 2001), the deception condition consisted of asking people to lie in response to yes/no questions by pressing one of two buttons. During any given series of trials the participants alternated between lying and telling the truth, depending on the color of a probe. The main finding of the study was bilateral activation in ventrolateral prefrontal cortex (BA 47) in the lie compared with the truth condition, a finding that was interpreted to reflect motor response inhibition. According to our taxonomy, this deception condition can be characterized as spontaneous and not fitting into a coherent scenario, and thus is a SI lie. Although we found activation in the anterior cingulate for SI lies [close to the medial frontal regions reported by Spence *et al.*. (2001)], which we interpreted as related to response inhibition, we did not find significant activation in BA 47 *per se* in the SI lie condition. This could be due to our random-effect analysis, which, while giving us more confidence in the generalizability of the results, was more conservative than the fixed-effect analysis reported by Spence *et al*. (Spence *et al*., 2001). Consistent with this hypothesis, we found a cluster of nine voxels for which the SI > T contrast was significant at $P < 0.005$ in right BA 47 when we compared SI lies with Truth in a spherical ROI (8 mm radius) centered at the coordinates reported in Spence *et al*. (Spence *et al*., 2001). It is also possible that alternating between lying and telling the truth in the same series of trials changes the strategies participants use to perform the task (Dove *et al*., 2000), and that BA 47 would not be more active if telling the truth and lying were carried out in separate sets of trials.

A comparison with the results reported by Lee *et al*. (Lee *et al*., 2002) and by Langleben *et al*. (Langleben *et al*., 2002) is more difficult because of major differences between the paradigms. However, the prefrontal–parietal network reported by Lee *et al*. (Lee *et al*., 2002) when people deliberately performed poorly on a set of arithmetic problems, or faked poor memory on a set of simple autobiographic questions, is generally consistent with our findings. The anterior cingulate activation found by Langleben *et al*. (Langleben *et al*., 2002) when people lied compared with when they told the truth in a simplified version of the GKT is also consistent with our findings.

All the fMRI studies of deception conducted so far, including ours, have used group analyses to detect difference between telling lies and telling the truth. This is a reasonable first step, but whether fMRI can become a useful tool for the detection of deception (setting aside for now important practical issues such as its cost) depends on whether reliable neural signatures of deception can be identified in single participants and in single trials. Thus, a substantial amount of research both on the deception paradigms and on the analysis methods remains to be conducted before we can fully assess the potential of fMRI as a lie detection device.

In summary, this study is a first attempt to demonstrate that different types of lies are associated with different patterns of brain activation. The results suggest that future neuroimaging studies of deception in more realistic settings should not collapse qualitatively different types of lies into a single category. This inappropriate pooling of results would increase variability in the data, and would obscure the ability to identify signals associated with specific types of deception.

To conclude, we must note that our two dimensions for characterizing types of lies are just the beginning. For example, lies are associated with a greater or lesser emotional response. A major limitation of our study compared to real settings is that the participants were not as emotionally involved in lying as they would be in a non-laboratory situation. Although we tried to use interesting scenarios about memorable events of their lives, the participants probably did not feel bad or guilty about lying (in fact, quite the reverse – they were cooperating by following the instructions – but were still engaged in prevarication). In addition, one can lie about one's own actions or about something one merely observed. We suspect that this dimension of self-involvement would also affect the types of processes that underlie the generation of lies. Thus, we are faced with at least four dimensions along which lies may vary, and there are undoubtedly more. An accurate and precise lie detection system will likely have to account for these dimensions and exploit the variations that arise from the different types of processing involved in the different types of lies.

Notes

We are grateful to Dr Haline E. Schendan for insightful comments and to Anne Smith at the McLean Hospital for technical help. This work was supported by a contract from the Office for Research and Development (Contract no. 98-F134600-000).

Address correspondence to Giorgio Ganis, Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge, MA 02138, USA. Email: ganis@wjh.harvard.edu.

References

Allen JJ, Iacono WG (1997) A comparison of methods for the analysis of event-related potentials in deception detection. Psychophysiology 34:234–240.

Andreasen NC, O'Leary DS, Paradiso S, Cizadlo T, Arndt S, Watkins GL, Ponto LL, Hichwa RD (1999) The cerebellum plays a role in conscious episodic memory retrieval. Hum Brain Mapp 8:226–234.

Birn RM, Cox RW, Bandettini PA (2002) Detection versus estimation in event-related fMRI: choosing the optimal stimulus timing. Neuroimage 15:252–264.

Braver TS, Barch DM, Gray JR, Molfese DL, Snyder A (2001) Anterior cingulate cortex and response conflict: effects of frequency, inhibition and errors. Cereb Cortex 11:825–836.

Buckner RL, Koutstaal W, Schacter DL, Wagner AD, Rosen BR (1998) Functional–anatomic study of episodic retrieval using fMRI. I. Retrieval effort versus retrieval success. Neuroimage 7:151–162.

Bunge SA, Ochsner KN, Desmond JE, Glover GH, Gabrieli JDE (2001) Prefrontal regions involved in keeping information in and out of mind. Brain 124:2074–2086.

Burgoon JK, Buller DB (1994) Interpersonal deception: III. Effects of deceit on perceived communication and nonverbal behavior dynamics. J Nonverb Behav 18:155–184.

Christoff K, Prabhakaran V, Dorfman J, Zhao Z, Kroger JK, Holyoak KJ, Gabrieli JDE (2001) Rostrolateral prefrontal cortex involvement in relational integration during reasoning. Neuroimage 14:1136–1149.

Cox RW (1996) AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput Biomed Res 29:162–173.

Cox R, Jesmanowicz A (1999) Real-time registration for functional fMRI. Magn Reson Med, 42:1014–1018.

DeSoto MC, Fabiani M, Geary DC, Gratton G (2001) When in doubt, do it both ways: brain evidence of the simultaneous activation of conflicting motor responses in a spatial Stroop task. J Cogn Neurosci 13:523–536.

Dove A, Pollmann S, Schubert T, Wiggins CJ, von Cramon DY (2000) Prefrontal cortex activation in task switching: an event-related fMRI study. Brain Res Cogn Brain Res 9:103–109.

Duncan J, Owen AM (2000) Common regions of the human frontal lobe recruited by diverse cognitive demands. Trends Neurosci 23:475–483.

Ekman P (1992) Telling lies. New York: WW Norton.

Ekman P (2001) Telling lies: clues to deceit in the marketplace, politics, and marriage. New York: WW Norton.

Epstein R, Harris A, Stanley D, Kanwisher N (1999) The parahippocampal place area: recognition, navigation, or encoding? Neuron 23:115–125.

Fletcher PC, Henson RN (2001) Frontal lobes and human memory: insights from functional neuroimaging. Brain 124:849–881.

Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC (1995) Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. Magn Reson Med 33:636–647.

Frank MG, Ekman P (1997) The ability to detect deceit generalizes across different types of high-stake lies. J Pers Soc Psychol 72:1429–1439.

Grady CL (1999) Neuroimaging and activation of the frontal lobes. In: The human frontal lobes: functions and disorders (Miller BL and Cummings JL, eds), pp. 196–230. New York: Guilford Press.

Horowitz SW, Kircher JC, Honts CR, Raskin DC (1997) The role of comparison questions in physiological detection of deception. Psychophysiology 34:108–115.

Horvath F, Jayne B, Buckley J (1994) Differentiation of truthful and deceptive criminal suspects in behavior analysis interviews. J Forensic Sci 39:793–807.

Kosslyn SM, Jolicoeur P (1980) A theory-based approach to the study of individual differences in mental imagery. In: Aptitude, learning, and instruction: cognitive processes analyses (Snow RE, Fredrico P-A, Montague WE, eds). Hillsdale, NJ: Erlbaum.

Kosslyn SM, Ganis G, Thompson WL (2001) Neural foundations of imagery. Nat Rev Neurosci 2:635–642.

Krause BJ, Schmidt D, Mottaghy FM, Taylor J, Halsband U, Herzog H, Tellmann L, Muller-Gartner HW (1999) Episodic retrieval activates the precuneus irrespective of the imagery content of word pair associates. A PET study. Brain 122:255–263.

Lane RD, Fink GR, Chau PM, Dolan RJ (1997) Neural activation during selective attention to subjective emotional responses. Neuroreport 8:3969–3972.

Langleben DD, Schroeder L, Maldjian JA, Gur RC, McDonald S, Ragland JD, O'Brien CP, Childress AR (2002) Brain activity during simulated deception: an event-related functional magnetic resonance study. Neuroimage 15:727–732.

Lee TMC, Liu H-L, Tan L-H, Chan CCH, Mahankali S, Feng C-M, Hou J, Fox PT, Gao J-H (2002) Lie detection by functional magnetic resonance imaging. Hum Brain Mapp 15:157–164.

Lykken DT (1974) Psychology and the lie detector industry. Am Psychol 29:725–739.

MacLaren VV (2001) A quantitative review of the guilty knowledge test. J Appl Psychol 86:674–683.

MacLeod AK, Buckner RL, Miezin FM, Petersen SE, Raichle ME (1998) Right anterior prefrontal cortex activation during semantic monitoring and working memory. Neuroimage 7:41–48.

Mehrabian A (1971) Nonverbal betrayal of feeling. J Exp Res Personality 5:64–73.

Office of Technology Assessment (1983) Scientific validity of polygraph testing: a research review and evaluation – a technical memorandum. US Office of Technology Assessment, Washington, DC.

Paivio A (1971) Image and verbal processes. New York: Holt, Rinehart, and Winston.

Podlesny JA, Raskin DC (1977) Physiological measures and detection of deception. Psychol Bull 84:782–799.

Raskin DC, Hare RD (1978) Psychopathy and detection of deception in a prison population. Psychophysiology 15:126–136.

Ruff CC, Woodward TS, Laurens KR, Liddle PF (2001) The role of the anterior cingulate cortex in conflict processing: evidence from reverse stroop interference. Neuroimage 14:1150–1158.

Schacter D (1996) Searching for memory. New York: Basic Books.

Smith EE, Jonides J (1998) Neuroimaging analyses of human working memory. Proc Natl Acad Sci USA 95:12061–12068.

Spence SA, Farrow TF, Herford AE, Wilkinson ID, Zheng Y, Woodruff PW (2001) Behavioural and functional anatomical correlates of deception in humans. Neuroreport 12:2849–2853.

Talaraich J, Tournoux P (1988) A co-planar stereotactic atlas of the human brain. Stuttgart: Thieme Verlag.

Vrij A (1994) The impact of information and setting on detection of deception by police detectives. J Nonverbal Behav 18:117–136.

Wagner AD, Desmond JE, Glover GH, Gabrieli JDE (1998) Prefrontal cortex and recognition memory. Functional-MRI evidence for context-dependent retrieval processes. Brain 121:1985–2002.

Wheeler ME, Petersen SE, Buckner RL (2000) Memory's echo: vivid remembering reactivates sensory-specific cortex. Proc Natl Acad Sci USA. 97:11125–9.

Xiong J, Gao J-H, Lancaster JL, Fox PT (1995) Clustered analysis for functional MRI activation studies of the human brain. Hum Brain Mapp 3:287–301.

Zuckerman M, DeFrank RS, Hall JA, Larrance DT, Rosenthal R (1979) Facial and vocal cues of deception and honesty. J Exp Soc Psychol 18:378–396.