# fMRI-activation patterns in the detection of concealed information rely on memory-related effects

Matthias Gamer,[1] Olga Klimecki,[2] Thomas Bauermann,[3] Peter Stoeter,[3] and Gerhard Vossel[4]

[1]Department of Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Germany, [2]Laboratory for Social and Neural Systems Research, University of Zurich, Switzerland, [3]Department of Neuroradiology, and [4]Department of Psychology, Interdisciplinary Research Group Forensic Psychophysiology, Johannes Gutenberg-University Mainz, Germany

**Recent research on potential applications of fMRI in the detection of concealed knowledge primarily ascribed the reported differences in hemodynamic response patterns to deception. This interpretation is challenged by the results of the present study. Participants were required to memorize probe and target items (a banknote and a playing card, each). Subsequently, these items were repeatedly presented along with eight irrelevant items in a modified Guilty Knowledge Test design and participants were instructed to simply acknowledge item presentation by pressing one button after each stimulus. Despite the absence of response monitoring demands and thus overt response conflicts, the experiment revealed a differential physiological response pattern as a function of item type. First, probes elicited the largest skin conductance responses. Second, differential hemodynamic responses were observed in bilateral inferior frontal regions, the right supramarginal gyrus and the supplementary motor area as a function of item type. Probes and targets were accompanied by a larger signal increase than irrelevant items in these regions. Moreover, the responses to probes differed substantially from targets. The observed neural response pattern seems to rely on retrieval processes that depend on the depth of processing in the encoding situation.**

**Keywords:** deception; guilty knowledge test; concealed information; skin conductance; response conflict

## INTRODUCTION

In the past few years, different experiments have been carried out in order to investigate the neural correlates of deception and information concealment. Whereas traditional forensic polygraphy relies on peripheral autonomic measures such as skin conductance, respiration and heart rate (e.g. Nakayama, 2002; Gamer *et al.*, 2006), recent research has focused on central nervous correlates of deception and concealed knowledge by using event-related brain potentials (ERPs, e.g. Rosenfeld *et al.*, 1988; Allen *et al.*, 1992; Rosenfeld *et al.*, 2004, 2006), Positron Emission Tomography (PET, Abe *et al.*, 2006) and fMRI (e.g. Langleben *et al.*, 2005; Phan *et al.*, 2005). Several experiments utilizing brain imaging techniques focused on intraindividual comparisons of deceptive and truthful responses (e.g. Spence *et al.*, 2001; Nuñez *et al.*, 2005), thus employing variations of the differentiation of deception paradigm (Furedy *et al.*, 1988, 1994; Gödert *et al.*, 2001).

However, one should bear in mind Lykken's notion (1998, p. 63ff.) in which he seriously questioned the existence of specific lie responses and therefore developed the Guilty Knowledge Test (GKT) (Lykken, 1959). Today, the GKT is

regarded as a scientifically controlled diagnostic test for the presence of information (Ben-Shakhar and Elaad, 2002) which is used in forensic investigations in Japan (Nakayama, 2002). The method is not directly based on deception (Furedy, 1986), but relies on intra-individual differences of the physiological responses to significant and insignificant items. Typically, peripheral autonomic reactions are recorded while the examinee is confronted with a number of multiple-choice questions asking for specific crime-related knowledge. Each question contains one crime-related detail, the probe item, and several irrelevant alternatives (e.g. What object was stolen last night? Was it (i) a ring, (ii) a purse, (iii) a notebook, (iv) a wristwatch, (v) a handbag?). The central assumption of the GKT is that persons with crime-related knowledge will exhibit stronger physiological reactions to probes than to irrelevant items due to the greater significance of probes and their storage in the episodic memory of the culprit.

Among the several neuroimaging studies which have investigated detection of deception in recent years, none has applied a strict GKT design as described above. Instead, most researchers used a combination of the differentiation of deception approach and the GKT (Langleben *et al.*, 2002, 2005; Kozel *et al.*, 2004a, b, 2005; Davatzikos *et al.*, 2005; Phan *et al.*, 2005). In these studies, subjects were typically instructed to truthfully acknowledge familiarity of one item while deceptively denying knowledge of another

item. By comparing the activity between deceptively and truthfully answered trials, the neural correlates of deceptive behavior were investigated. These mainly included the anterior cingulate cortex (ACC) and the inferior and middle frontal gyrus (Langleben *et al.*, 2002, 2005; Kozel *et al.*, 2004a, b, 2005; Phan *et al.*, 2005). Studies which directly employed the differentiation of deception paradigm consistently reported a stronger activation in the ventrolateral prefrontal cortex with a slight dominance of the right side, medial prefrontal regions and the ACC when deceptive responses were contrasted with truthful responses (Spence *et al.*, 2001; Nuñez *et al.*, 2005; Abe *et al.*, 2006). Additionally, significant activations of premotor or motor areas during deception were described (e.g. Spence *et al.*, 2001; Langleben *et al.*, 2002, 2005).

Although activation patterns differed substantially between studies (which might be due to differences in the particular research paradigms), almost all studies identified the contribution of the frontal cortex to deceptive responding. This finding suggests a central role of working memory, response monitoring and attentional processes. The designs described, however, do not permit a dissociation between processes related to deception, response conflict and memory retrieval. Moreover, these neural activation patterns have been rarely related to the peripheral autonomic reactions used in traditional polygraphy.

In a recent study, however, Gamer *et al.* (2007) provided a first link between behavioral measures, sympathetic arousal and neural activation patterns during a GKT examination. In a three item GKT design (cf., Farwell and Donchin, 1991), participants were instructed to conceal knowledge of specific probe items while viewing a randomized sequence of probes and irrelevant items. Additionally, rare targets were occasionally presented which required a different behavioral response. It turned out that rare and significant items (probes and targets) were accompanied by an increase of activity in inferior frontal and mid-cingulate regions when compared to irrelevant items. Target presentation specifically involved a stronger activation of contralateral motor and somatosensory regions (cf., Spence *et al.*, 2001; Langleben *et al.*, 2002, 2005). The differential activation reported for the right inferior frontal region was modulated by stimulus conflicts that were also reflected in the behavioral data. Furthermore, activity in this region was significantly correlated with SCR amplitudes and response times as revealed by parametric analyses. These results suggest that processes related to attention and memory play a central role in evoking distinct hemodynamic response patterns for different item types in a GKT examination. However, since all stimuli required response monitoring, it was not possible to directly assess the relative contribution of memory related processes and response conflicts to the physiological activation pattern. Therefore, the current GKT study, which closely followed the experimental procedure of Gamer *et al.* (2007), aimed at investigating whether the

omission of overt reaction conflicts alters neural activation patterns and differential skin conductance responding. Although reaction conflicts and deceptive responses are usually part of a GKT examination (Ben-Shakhar and Elaad, 2003), it has been repeatedly reported that even in the absence of any verbal or motor response, autonomic reactions can differ significantly between probe and irrelevant items when the examinee recognizes the relevant details (e.g. Janisse and Bradley, 1980; Horneman and O'Gorman, 1985; Elaad and Ben-Shakhar, 1989; Bradley *et al.*, 1996). These results substantiate Lykken's original claim (1959) that the physiological response pattern obtained in GKT examinations mainly relies on recognition memory and not on deception.

Based on this reasoning, we expected to find a differential physiological response pattern in skin conductance and fMRI data for probes and irrelevant items which should even occur in the absence of overt response conflicts. Therefore, in the current GKT study, participants were simply instructed to confirm the presentation of each item, so that neither deception nor conflicts between competing behavioral responses were involved.

With regard to the fMRI data, we first assumed that there would be no differential activity in parietal somatosensory and motor regions, since our study did not involve differential behavioral reactions. Secondly, we expected to find lateralized inferior frontal activation differences in both hemispheres as a function of item type due to effects of memory retrieval (Iidaka *et al.*, 2006) and target detection (Linden *et al.*, 1999). Finally, we hypothesized that no differences in skin conductance responses or neural activation patterns would occur between the two stimulus categories (playing cards and bank notes) because in contrast to Gamer *et al.* (2007), there was no need for inhibiting predominant but wrong responses in the current study.

## METHOD
### Participants
Twenty-eight right-handed subjects with no report of neurological or psychiatric disorders volunteered to participate in the study and gave written informed consent according to the Declaration of Helsinki. Due to technical problems, the data of five subjects could not be included in the analyses. The final sample consisted of 23 persons, mainly students, of whom five were female and 18 male. Their mean age was 25.1 years (s.d. = 4.6 years), with a range of 20–42 years.

### Instruments
Skin conductance and behavioral data (reaction times and accuracy) were recorded continuously and saved for offline analysis. Skin conductance was measured using a constant voltage system (0.5 V) with two Ag/AgCl electrodes (0.8 cm diameter) that were filled with 0.05 M NaCl electrolyte. The electrodes were attached to the skin surface at the medial side

of the right foot (Edelberg, 1967). Skin conductance was recorded at 10 Hz by a conventional personal computer. Reaction times were acquired using a fiberoptic response keypad system (LumiTouch; Photon Control Inc., Vancouver, Canada).

## Design

A $2 \times 3$ fully crossed within-subject design with the factors stimulus category (playing cards *vs* bank notes) and item type (probe, target, irrelevant) was used in the experiment. It was derived from research on electrophysiological correlates in the detection of concealed information (e.g. Farwell and Donchin, 1991) and closely matched the experimental design of a recent fMRI study on the GKT (Gamer *et al.*, 2007). In essence, subjects were presented with three item types, which differed with respect to their personal significance. *Probe items* were represented by a jack of spades and a 20-euro bank note. Knowledge of these items was acquired by the participant prior to the physiological measurements and participants were instructed not to reveal this knowledge until the end of the experiment. A king of spades and a 100-euro bank note served as *targets*. Subjects saw them prior to the physiological measurements and had to memorize them. *Irrelevant Items* (the playing cards 9, 10, queen and ace of spades and the bank notes 5, 10, 50, 200 euros) were not presented separately and thus had no specific meaning in the experimental context. We decided to use playing cards and bank notes as stimulus categories because previous research has shown that the higher perceptual similarity between probe and target item in the category of the playing cards enhanced response conflicts and modulated neural activity in the lateral prefrontal cortex (Gamer *et al.*, 2007). We were specifically interested in whether this response pattern would also hold for a different behavioral task that did not require a precise response monitoring. During the measurement of physiological responses, each stimulus was displayed 20 times in an event-related design. This resulted in a total presentation of 40 probes, 40 targets and 160 irrelevant stimuli.

## Procedure

Participants had to choose one of three envelopes not knowing that all three envelopes contained the same playing card and bank note. Subjects were instructed to memorize the content of the chosen envelope (probes) and to hide the items in their pocket. Subjects were told not to reveal knowledge of these probe items until the end of the experiment. After attaching the skin conductance electrodes, the experimenter presented the target items to the participants with the instruction to memorize them. To ensure that participants paid attention to the stimulus presentation, they were required to press a key with their right index finger as fast as possible each time a stimulus appeared. Furthermore, participants were forewarned that they would have to accomplish a memory test after fMRI scanning. This

was done to ensure that participants were actually paying attention to each stimulus instead of stereotypically pressing the response button whenever a screen change occurred.

Stimuli were rear-projected on the window between scanning and control room and could be viewed with a $45°$ angled mirror attached to the head coil. The viewing angle approximately subtended $5-8°$ in the horizontal and vertical direction. Stimuli were presented in a pseudo randomized sequence which was identical for each participant. The first four stimuli were not used for further analyses and showed irrelevant items. The remaining sequence consisted of four blocks, each having the same amount of probes, targets and irrelevants. Each stimulus was presented for 1 s; a fixation cross was displayed between stimuli. The interstimulus interval lasted 5.2, 8.3, 11.4, 14.5 or 18.6 s, with an additional random jittering of 0–1 s relative to scan onset. This resulted in effective interstimulus intervals of 5.2–19.2 s ($M = 7.7$ s, s.d. $= 2.1$ s). After the measurement session, each participant completed a memory test about the items presented in the experiment. All participants were able to reproduce the items that were displayed in the measurement session. Most importantly, probe and target items could be specified correctly by all participants.

## Image acquisition

Blood oxygen level dependent (BOLD) contrast signals were acquired with a clinical whole-body MRI scanner and a standard polarized head coil (Magnetom Vision, Siemens Erlangen, Germany) at a magnetic field strength of 1.5 Tesla. First, isotropic high resolution ($1 \times 1 \times 1 \text{ mm}^3$) structural images were recorded using a T1-weighted sagittal oriented mprage sequence with 180 slices. $T_2{}^*$-weighted images were acquired using an echo planar imaging (EPI) pulse sequence (TR $= 3100$ ms, TE $= 60$ ms, FOV $= 192 \times 192 \text{ mm}^2$, th $= 5$ mm without gap, voxel size $= 3 \times 3 \times 5 \text{ mm}^3$). The protocol included 610 volumes, each consisting of 28 transversal slices. The volumes were oriented parallel to the anterior–posterior commissural plane and covered the entire brain of each subject. The first three volumes were discarded to allow for signal equilibration.

## Data preprocessing and analysis

*Behavioral data.*  To ensure that participants paid attention to the stimulus presentations, the proportion of correctly confirmed stimulus presentations and the mean reaction times were computed. The proportion of confirmed stimulus presentations was 99.3% with a mean reaction time of 448 ms (s.d. $= 129$ ms). None of these behavioral measures differed as a function of stimulus category or item type.

*Skin conductance responses.*  Amplitudes of skin conductance responses that began between 1 and 3 s after stimulus onset were scored as stimulus-evoked responses and measured as changes in micro-Siemens if they exceeded a threshold of 0.01 μS. When overlapping responses were observed (1.9% of all trials), the inflection point between

the two responses served as the baseline or peak, depending on the latency criterion. The amplitudes were log-transformed according to a formula by Venables and Christie (1980). Trials with missing behavioral responses were excluded from the SCR analysis. In order to test for a differential habituation of the electrodermal responses, SCR amplitudes were examined by dividing the total number of trials into four temporal blocks, each containing the same amount of probes, targets and irrelevant items.

To circumvent the problematic sphericity assumption of univariate repeated measures analyses of variance, skin conductance data were analyzed using multivariate analyses of variance (MANOVAs) with the within subjects factors temporal block, stimulus category and item type (Vasey and Thayer, 1987).

*Imaging data.* Statistical Parametric Mapping (SPM5, Wellcome Department of Imaging Neuroscience, London, UK) was used for preprocessing and analyzing the imaging data. In a first step, each participant's data were slice time-corrected and motion-corrected. Functional data were then coregistered with the corresponding high resolution T1 image. The T1 images were spatially normalized to the standard anatomical space (T1 MNI template) and transformation parameters were subsequently applied to all functional images within each participant. Functional images were smoothed with an $8 \times 8 \times 12\,mm^3$ full width at half maximum Gaussian kernel and high pass filtered at 128 s. The correction for autocorrelation [AR(1) method] between scans was applied.

Differential activation as a function of item type (probe, target and irrelevant) was examined using a general linear model (GLM) analysis on an individual level. The three item types were modeled as separate regressors and then convolved with the hemodynamic response function. The first four (irrelevant) items and trials with missing behavioral responses were modeled separately as covariates of no interest. Additionally, the six movement parameters that were estimated during preprocessing were included in the design matrix. Simple contrast maps (each item type *vs* baseline condition) were generated for each subject individually and then entered into a random-effects analysis (one-way ANOVA within subjects) to identify regions which showed significant activation differences between item types (*F*-contrast). The relative contribution of both stimulus categories to this main effect was examined by post hoc analyses on the average percentage signal change in regions of interest (ROIs) that were functionally defined by the *F*-contrast of the random effects analysis described above (Gamer *et al.*, 2007). These values were obtained using the SPM-toolbox MarsBaR (Brett *et al.*, 2002; http://marsbar.sourceforge.net).

Since the current study was highly similar to the study of Gamer *et al.* (2007), it was possible to directly compare the results of both studies statistically in order to examine the relevance of overt response conflicts for the neural activation pattern. To this aim, simple contrast maps reflecting the activation of each item type *vs* baseline condition were generated for each study and each subject individually and then entered into a two-way random-effects analysis (ANOVA) using the group factor study (the current study *vs* the data of Gamer *et al.*, 2007) and the within-subject factor item type. To identify regions that were consistently more activated following the presentation of probe items as compared to irrelevants in both studies, we first calculated a conjunction analysis (Nichols *et al.*, 2005). In a second step, an interaction analysis was carried out to reveal differential activations of probe and irrelevant items in both studies.

For all random-effects SPM-analyses, *P*-values were corrected for multiple comparisons using a false discovery rate (FDR) of 0.05 (Genovese *et al.*, 2002). Additionally, activations were required to reach a spatial extent threshold of at least 20 contiguous voxels.

## RESULTS

### Skin conductance responses

A $4 \times 2 \times 3$ MANOVA on the log-transformed SCR amplitudes using the temporal block, the stimulus category and the item type as within-subject factors revealed statistically significant main effects of temporal block, $F(3, 20) = 9.30$, $P < 0.001$, Cohen's $f = 0.29$, and item type, $F(2, 21) = 3.88$, $P < 0.05$, $f = 0.09$. All other effects were insignificant. As can be seen from Figure 1, SCRs decreased across the experiment which indicates that habituation occurred. Post hoc comparisons, using the Scheffé method, revealed larger SCR responses to probes ($M = 0.08$; s.d. $= 0.08$) than to irrelevants ($M = 0.06$; s.d. $= 0.07$). Furthermore, SCRs to targets ($M = 0.07$; s.d. $= 0.06$) tended to differ from probes ($P < 0.1$) but not from irrelevant items.

### Imaging results

Significant activation differences as a function of item type were observed bilaterally in inferior frontal regions incorporating parts of the insula, in the right supramarginal gyrus and in the right supplementary motor area (Table 1, Figure 2A). Separate $2 \times 3$ MANOVAs on the percentage signal increase in each functionally defined ROI were calculated to clarify the influence of the within subject factors stimulus category (playing card *vs* bank note) and item type (probe, target, or irrelevant) on the main effect described above. The main effect of item type was significant for all ROIs, but, most importantly, neither a main effect of stimulus category nor an interaction of stimulus category and item type reached statistical significance in any ROI. Thus, differences in hemodynamic responses between item types did not change as a function of stimulus category. As can be seen from Figure 3, probes were consistently accompanied by a larger percentage signal change than irrelevant items in all ROIs. Furthermore, the percentage signal change elicited by targets was positioned in between
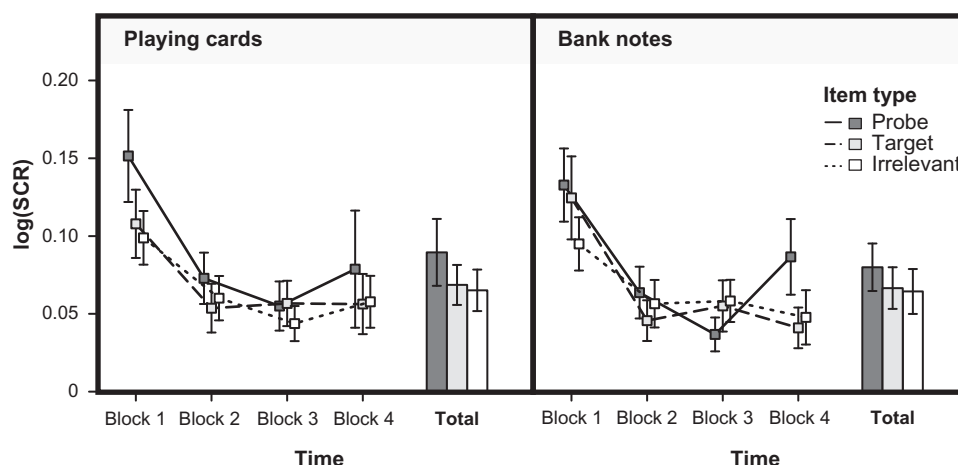
**Fig. 1** Temporal course of the log-transformed skin conductance response (SCR) amplitudes as a function of stimulus category (playing cards *vs* bank notes) and item type. Additionally, mean values across temporal blocks are displayed. Error bars indicate standard errors of the mean.

**Table 1** Regional brain activity differing between item types

| Regions contained within cluster | Peak voxel (MNI coordinates) | | | Cluster size | | |
|---|---|---|---|---|---|---|
| | x | y | z | (voxels) | $F(2, 44)$ | Corrected $P$ |
| **Right inferior frontal gyrus (p. Opercularis)** | 48 | 18 | 9 | 261 | 18.97 | 0.011 |
| Right insula | 45 | 15 | −3 | | 17.87 | 0.011 |
| Right inferior frontal gyrus (p. Triangularis) | 54 | 21 | 6 | | 16.58 | 0.011 |
| Right inferior frontal gyrus (p. Opercularis) | 39 | 12 | 30 | | 12.44 | 0.018 |
| **Left insula** | −36 | 21 | −3 | 81 | 18.19 | 0.011 |
| Left inferior frontal gyrus (p. Triangularis) | −42 | 27 | 3 | | 16.05 | 0.011 |
| **Right supramarginal gyrus** | 57 | −48 | 24 | 36 | 13.60 | 0.014 |
| Right supramarginal gyrus | 63 | −36 | 30 | | 10.91 | 0.029 |
| **Right supplementary motor area** | 12 | 6 | 69 | 23 | 13.60 | 0.014 |
| Right supplementary motor area | 6 | 12 | 60 | | 11.45 | 0.024 |

*Note*: Regions included in a single cluster are listed together. Regions containing the peak voxel are printed in bold. The spatial extent of each cluster was ≥ 20 voxel and an FDR of 0.05 was used to correct for multiple comparisons.

the responses to probes and irrelevant items in all ROIs except the left inferior frontal gyrus. In this region, probes elicited larger responses than targets and irrelevant items, whereas the latter two did not differ significantly.

The conjunction analysis of the current data in comparison to Gamer *et al.* (2007) revealed that probe items elicited a larger activation of bilateral inferior frontal regions including the right insula as compared to irrelevants in both studies (Table 2, Figure 2B). In contrast, there was no region that showed a differential response to the contrast of probes *vs* irrelevants when the two studies were compared. Only when lowering the threshold to $P < 0.001$ (uncorrected), a small region in the right inferior frontal gyrus (p. Orbitalis, peak voxel = 45, 24, −15, $t[70] = 4.04$) came up that showed a larger activation difference in the study of Gamer *et al.* (2007) as compared to the current study (Figure 2B).

## DISCUSSION

The central aim of this study was to differentiate neural processes related to response monitoring from processes

of memory retrieval in the GKT. Most previous neuroimaging studies on the detection of concealed knowledge (e.g. Langleben *et al.*, 2002, 2005; Phan *et al.*, 2005) focused on deception and did not pay much attention to memory-related effects, although the original concept of the GKT (Lykken, 1959) strongly relies on recognition memory. With the aim of studying the elementary neural processes involved in GKT-like paradigms, we decided to use a very restricted experimental setting. We eliminated response selection demands and overt reaction conflicts to assure that item types differed only with respect to their personal significance. In addition, we measured skin conductance responses during fMRI data acquisition, as SCRs are generally accepted to be the most reliable indicator of the orienting response (Barry, 1984), which in turn is assumed to be the key mechanism for triggering differential autonomic reactions in traditional GKT examinations (Verschuere *et al.*, 2004).

Corresponding to other research on the GKT, probes elicited larger SCRs than irrelevant items (Ben-Shakhar
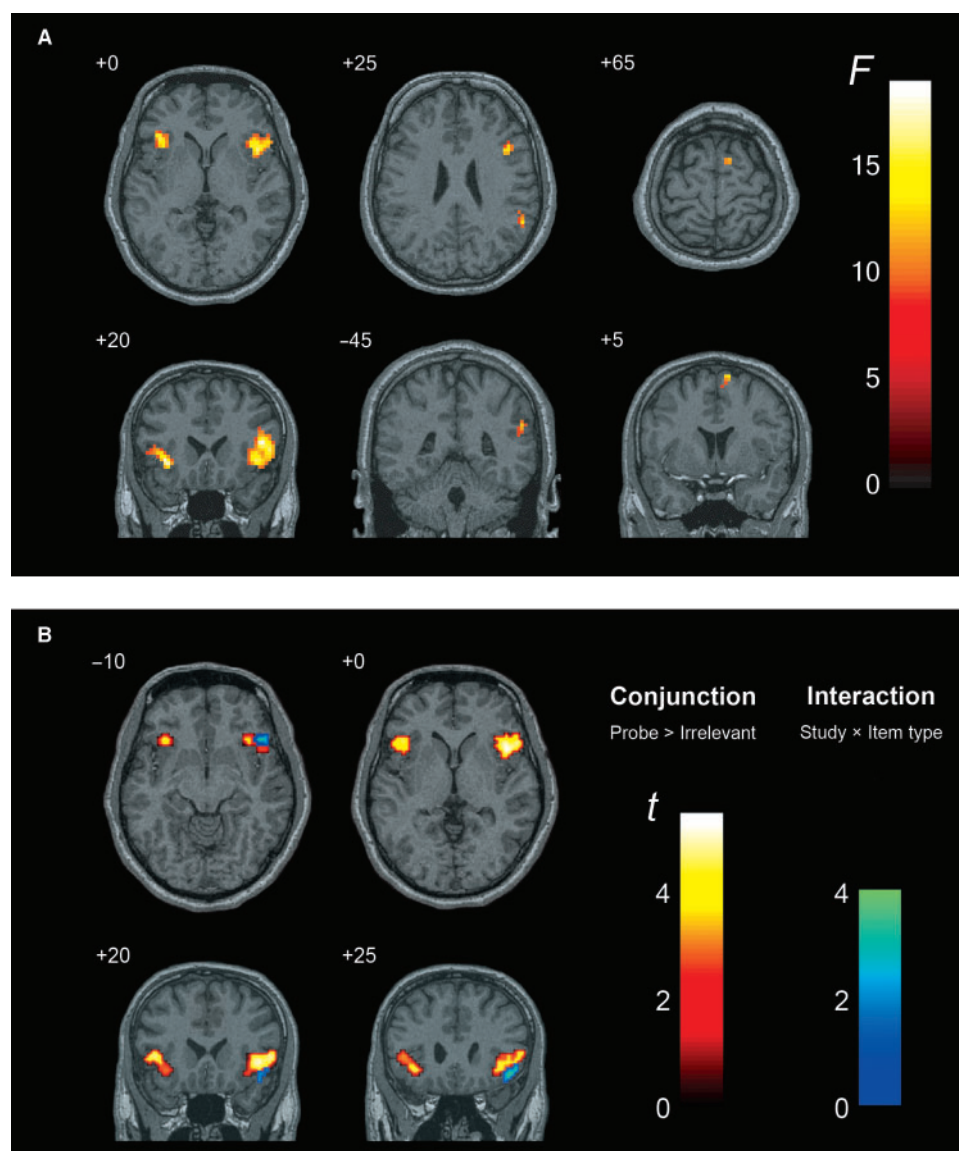
**Fig. 2** Brain areas showing a significant main effect of item type in the group analysis. Regions are displayed on axial and coronal slices of a single subject $T_1$-weighted structural image, which was spatially normalized to the standard anatomical space using the MNI template brain. Slice location is given by its respective MNI coordinate and images are shown in neurological convention (left side of image is left side of brain). Color scales represents $F$ or $t$ score values, respectively, for corresponding functional overlays. (**A**) Main effect of item type in a one-way analysis of variance within subjects. (**B**) Common and differential activations elicited by probes as compared to irrelevant items in the current study and Gamer *et al.* (2007).

and Elaad, 2003), thus, indicating an electrodermal differentiation between meaningful and irrelevant items. Furthermore, target items, which were just presented shortly prior to the onset of the experiment, tended to be associated with smaller SCR amplitudes than probes, although both item types were memorized before the experiment and both were presented with the same frequency during the experiment. This result suggests that, although no response selection was required, participants might have experienced an implicit response conflict since probe items carried the potential of being involuntarily identified as potentially relevant stimuli without requiring specific overt responses. Furthermore, probes were processed more deeply than targets (Craik and Lockhart, 1972), which might have

contributed to the differential electrodermal and hemodynamic response pattern.

With respect to neural activity, we expected that the somatosensory and motor regions that were previously shown to be related to deceptive responding (e.g. Spence *et al.*, 2001; Langleben *et al.*, 2002, 2005; Gamer *et al.*, 2007) should not be differentially activated as a function of item type since no response selection was required in our task. In fact, activity of the primary motor cortex and the adjacent parietal regions did not differ significantly between item types. We did, however, find a differential activation of the supplementary motor area (SMA) as a function of item type that could reflect an orienting response to significant stimuli (cf. Linden *et al.*, 1999; Downar *et al.*, 2000, 2001).
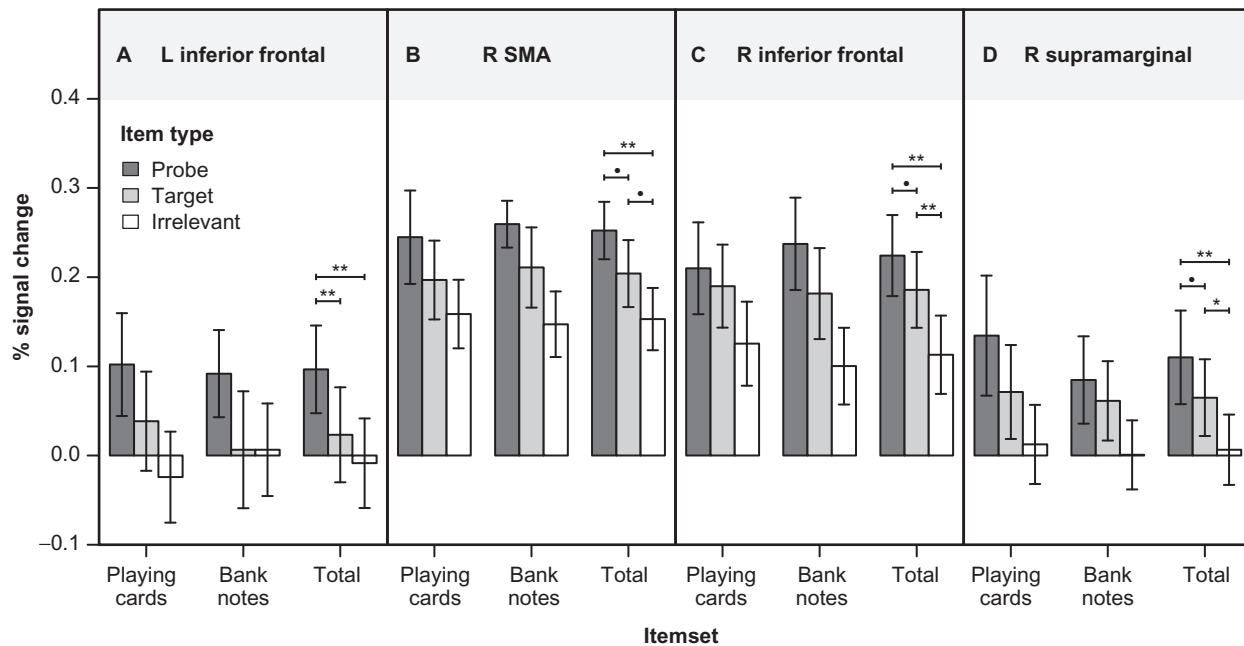
**Fig. 3** Mean percentage signal change in the left inferior frontal gyrus (**A**), the right supplementary motor area (SMA, **B**), the right inferior frontal gyrus (**C**) and the right supramarginal gyrus (**D**) as a function of stimulus category and item type. The regions of interest were functionally defined by the main effect of item type. Mean values across item categories are displayed on the right side of each panel. Error bars indicate standard errors of the mean. The main effect of item type was followed by pairwise post hoc comparisons, using the Scheffé method. Significance stars correspond to, •$P < 0.10$, *$P < 0.05$, **$P < 0.01$.

**Table 2** Brain regions that are commonly stronger activated by viewing probes compared to irrelevants in the current study and the study of Gamer *et al.* (2007)

| Regions contained within cluster | Peak voxel (MNI coordinates) | | | Cluster size | | |
|---|---|---|---|---|---|---|
| | $x$ | $y$ | $z$ | (voxels) | $t(70)$ | Corrected $P$ |
| **Right insula** | 45 | 18 | −3 | 172 | 5.48 | 0.006 |
| Right inferior frontal gyrus (p. Triangularis) | 54 | 21 | 3 | | 5.25 | 0.007 |
| **Left inferior frontal gyrus (p. Triangularis)** | −42 | 21 | 3 | 71 | 4.54 | 0.007 |
| Left inferior frontal gyrus (p. Orbitalis) | −36 | 24 | −9 | | 4.54 | 0.007 |

*Note*: Regions included in a single cluster are listed together. Regions containing the peak voxel are printed in bold. The spatial extent of each cluster was $\geq$ 20 voxel. An FDR of 0.05 was used for the conjunction analysis to correct for multiple comparisons.

Concerning the second hypothesis, we expected to find a stronger activation of lateral inferior frontal regions for recognized (probes and targets) as compared to irrelevant items. Indeed, the hemodynamic response pattern differed as a function of item type in ventrolateral and inferior frontal regions of both hemispheres. Comparable activation differences in the context of deception were reported by a large number of studies (Spence *et al.*, 2001; Lee *et al.*, 2002, 2005; Kozel *et al.*, 2004a, b, 2005; Davatzikos *et al.*, 2005; Langleben *et al.*, 2005; Nuñez *et al.*, 2005; Phan *et al.*, 2005; Abe *et al.*, 2006; Gamer *et al.*, 2007). This activation pattern was repeatedly interpreted as being related to overt response conflicts (Aron, 2004; Nuñez *et al.*, 2005; Gamer *et al.*, 2007). In the current study, however, such response conflicts did not occur since all items were associated with the same behavioral response.

Instead, two other explanations should be considered: First, lateral prefrontal activations were repeatedly found in a variety of different memory tests such as working memory tasks, episodic retrieval and semantic memory tests (Wagner *et al.*, 1998; D'Esposito *et al.*, 2000; Konishi *et al.*, 2000; Nyberg *et al.*, 2003; Ranganath *et al.*, 2003). With respect to the current study, probes and targets, which were both associated with a specific episodic experience, could be successfully differentiated from irrelevant items. This suggests that retrieval processes might be responsible for the differential activation of lateralized prefrontal regions as a function of item type.

Additionally, the presentation frequency of each item type should be considered as a relevant factor: taken together, items of personal significance (probes and targets) were shown less frequently (in total on one third of all trials)

than irrelevant items. Similar to our findings, Strange *et al.* (2000) reported a larger activation of lateral prefrontal areas for rare events, which indicates that enhanced activation of these regions might be related to a stimulus categorization process which takes into account the relative frequency of items from different stimulus classes (cf. Braver *et al.*, 2001). The response pattern observed in the current study, however, cannot be entirely explained by this account, because although probes and targets were presented with the same frequency, probes were consistently associated with a larger percentage signal increase than targets in all functionally defined ROIs. This effect can be explained, when taking into account that probes were encoded more deeply than targets, since the latter were only presented briefly before the beginning of the examination (cf. Craik and Lockhart, 1972). Recent studies have shown that a rare presentation of previously learned items (Herron *et al.*, 2004) as well as a deeper encoding (Iidaka *et al.*, 2006) are related to increased responses in lateral frontal regions. According to this reasoning, memory related effects are responsible for a differential activation of lateral prefrontal regions in a GKT examination with probes eliciting a larger response than targets because of their deeper encoding. Additionally, this response pattern might be amplified by the infrequent presentation of probes and targets as compared to irrelevant items. The conjunction analysis that was carried out to statistically compare the results of the current study to the data of Gamer *et al.* (2007) substantiates this reasoning by demonstrating that probes elicited a larger activation of inferior frontal regions irrespective of whether overt response conflicts were involved or not. The significant interaction of study and item type in the right inferior frontal cortex on a less stringent threshold may indicate that response conflicts can further enhance this differential activation (Aron *et al.*, 2004).

Another region that was differentially activated as a function of item type in our study was the right supramarginal gyrus. Again, several other researchers reported significant differences in hemodynamic responses between deceptively and truthfully answered trials in this area (Kozel *et al.*, 2005; Langleben *et al.*, 2005; Lee *et al.*, 2005). However, a more general explanation of these results should be considered. One suggestion is provided by several studies using an odd-ball paradigm that requires participants to detect rare targets in a train of frequent non-target stimuli (e.g. Linden *et al.*, 1999). In these studies, the supramarginal gyrus was shown to be consistently activated along with lateral prefrontal regions when contrasting the neural activation of targets and non-targets (e.g. MacCarthy, 1997; Kiehl *et al.*, 2001;). A similar response pattern is obtained when participants merely observe infrequent changes in stimulation without responding behaviorally to them (Downar *et al.*, 2000, 2002). Taken together, these studies suggest that the supramarginal gyrus plays an important role in identifying salient stimuli (cf., Downar *et al.*, 2001). These findings are in line

with the response pattern observed in the current study, in which probes and targets resembled the less frequent item category. Furthermore, the act of hiding the probe items in a pocket might have resulted in memory related effects which resemble a study by Russ *et al.* (2003) where a substantial increase of activity in the supramarginal gyrus was reported as a function of enactment.

In our third hypothesis, we postulated that no differences would be observed in activation patterns between stimulus categories since no response selection was required and therefore stimulus conflicts should be of no relevance. This assumption was substantiated by our results, which show that physiological responses differed significantly as a function of item type, whereas neither the skin conductance data nor the neural activation pattern varied between stimulus categories (i.e. playing cards *vs* bank notes).

Beside a very recent preliminary study that examined only three participants (Hakun *et al.*, 2008), this is the first neuroimaging study on the detection of concealed information which utilized a GKT-like paradigm without response selection. The design employed permitted us to directly examine neural processes that underlie the autonomic response pattern observed in the GKT and related techniques (Ben-Shakhar and Elaad, 2003). Our findings, which are largely comparable to the results of Hakun *et al.* (2008), suggest that deception can be ruled out as a necessary condition for evoking a distinct pattern of neural activity as a function of item type (see also Ben-Shakhar & Furedy, 1990, pp. 103f). The response pattern observed in the current study rather seems to rely on the automatic differentiation of rare and meaningful information (probes and targets) from irrelevant stimuli. Moreover, it was demonstrated that the physiological responses were sensitive to the encoding richness of retrieved information in terms of processing depth. Taken together, our results comply with other studies on the GKT that employed autonomic measures (Ben-Shakhar and Elaad, 2003) or ERPs (Farwell and Donchin, 1991; Allen *et al.*, 1992; van Hooff *et al.*, 1996; Farwell and Smith, 2001; Rosenfeld *et al.*, 1988, 2004, 2006).

With regard to potential forensic applications, our findings suggest that the differential physiological responses to probes and irrelevant items might be even larger in a forensic context under the assumption that the culprit processed the crime-relevant information deeply. This should be the case for items that were prepared before committing a crime or circumstances that were encountered during the course of the offense and directly affected the actions of the perpetrator (Nakayama, 2002). It should be noted, however, that standardized laboratory studies cannot be generalized to forensic field examinations because the emotional and motivational involvement of the examinee presumably differs largely between these contexts (Sip *et al.*, 2008). Moreover, certain populations (e.g. psychopaths) might exhibit abnormal response patterns due to differences in their brain structure (Raine *et al.*, 2000). Despite these

restrictions, the GKT is a widely accepted, standardized test that is based on scientific principles. Future studies, combining imaging techniques with autonomic measures might help to further elucidate the psychological and physiological substantiation of the GKT (National Research Council, 2003) as well as to clarify the utility of neuro-imaging techniques in this domain (Bles and Haynes, 2008). Importantly, a well-grounded theoretical understanding and a careful examination of social and ethical concerns are necessary before fMRI can be reasonably applied in forensic settings (cf., Wolpe *et al.*, 2005; Greely and Illes, 2007).

## REFERENCES

Abe, N., Suzuki, M., Tsukiura, T., et al. (2006). Dissociable roles of prefrontal and anterior cingulate cortices in deception. *Cerebral Cortex*, 16, 192–99.

Allen, J.J., Iacono, W.G., Danielson, K.D. (1992). The identification of concealed memories using the event-related potential and implicit behavioral measures: a methodology for prediction in the face of individual differences. *Psychophysiology*, 29, 504–22.

Aron, A.R., Robbins, T.W., Poldrack, R.A. (2004). Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences*, 8, 170–7.

Barry, R.J. (1984). Preliminary processes in O-R elicitation. *Acta Psychologica*, 55, 109–42.

Ben-Shakhar, G., Furedy, J.J. (1990). *Theories and applications in the detection of deception: A psychophysiological and international perspective.* New York: Springer.

Ben-Shakhar, G., Elaad, E. (2002). The Guilty Knowledge Test (GKT) as an application of psychophysiology: Future prospects and obstacles. In: Kleiner, M., editor. *Handbook of polygraph testing.* San Diego: Academic Press, pp. 87–102.

Ben-Shakhar, G., Elaad, E. (2003). The validity of psychophysiological detection of information with the Guilty Knowledge Test: a meta-analytic review. *Journal of Applied Psychology*, 88, 131–51.

Bles, M., Haynes, J. (2008). Detecting concealed information using brain-imaging technology. *Neurocase*, 14, 82–92.

Bradley, M.T., MacLaren, V.V., Carle, S.B. (1996). Deception and non-deception in guilty knowledge and guilty action polygraph tests. *Journal of Applied Psychology*, 81, 153–60.

Braver, T.S., Barch, D.M., Gray, J.R., Molfese, D.L., Snyder, A. (2001). Anterior cingulate cortex and response conflict: effects of frequency, inhibition and errors. *Cerebral Cortex*, 11, 825–36.

Brett, M., Anton, J., Valabregue, R., Poline, J. (2002) Region of interest analysis using an SPM toolbox [Abstract]. *Presented at the 8th International Conference on Functional Mapping of the Human Brain, June 2–6, 2002, Sendai, Japan.* Available on CD-ROM in *Neuroimage*, 16.

Craik, F.I.M., Lockhart, R.S. (1972). Levels of processing: a framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671–84.

D'Esposito, M., Postle, B.R., Rypma, B. (2000). Prefrontal cortical contributions to working memory: evidence from event-related fMRI studies. *Experimental Brain Research*, 133, 3–11.

Davatzikos, C., Ruparel, K., Fan, Y., et al. (2005). Classifying spatial patterns of brain activity with machine learning methods: application to lie detection. *Neuroimage*, 28, 663–8.

Downar, J., Crawley, A.P., Mikulis, D.J., Davis, K.D. (2000). A multimodal cortical network for the detection of changes in the sensory environment. *Nature Neuroscience*, 3, 277–83.

Downar, J., Crawley, A.P., Mikulis, D.J., Davis, K.D. (2001). The effect of task relevance on the cortical response to changes in visual and auditory stimuli: an event-related fMRI study. *Neuroimage*, 14, 1256–67.

Downar, J., Crawley, A.P., Mikulis, D.J., Davis, K.D. (2002). A cortical network sensitive to stimulus salience in a neutral behavioral context across multiple sensory modalities. *Journal of Neurophysiology*, 87, 615–20.

Edelberg, R. (1967). Electrical properties of the skin. In: Brown, C.C., editor. *Methods in Psychophysiology.* Baltimore: Williams and Wilkins, pp. 1–53.

Elaad, E., Ben-Shakhar, G. (1989). Effects of motivation and verbal-response type on psychophysiological detection of information. *Psychophysiology*, 26, 442–51.

Farwell, L.A., Donchin, E. (1991). The truth will out: interrogative polygraphy (lie detection) with event-related brain potentials. *Psychophysiology*, 28, 531–47.

Farwell, L.A., Smith, S.S. (2001). Using brain MERMER testing to detect knowledge despite efforts to conceal. *Journal of Forensic Sciences*, 46, 135–43.

Furedy, J.J. (1986). Lie detection as psychophysiological differentiation: Some fine lines. In: Coles, M., Donchin, E., Porges, S., editors. *Psychophysiology: Systems, Processes, and Applications – A handbook.* New York: Guilford, pp. 683–700.

Furedy, J.J., Davis, C., Gurevich, M. (1988). Differentiation of deception as a psychological process: a psychophysiological approach. *Psychophysiology*, 25, 683–8.

Furedy, J.J., Gigliotti, F., Ben-Shakhar, G. (1994). Electrodermal differentiation of deception: the effect of choice versus no choice of deceptive items. *International Journal of Psychophysiology*, 18, 13–22.

Gamer, M., Rill, H., Vossel, G., Gödert, H.W. (2006). Psychophysiological and vocal measures in the detection of guilty knowledge. *International Journal of Psychophysiology*, 60, 76–87.

Gamer, M., Bauermann, T., Stoeter, P., Vossel, G. (2007). Covariations among fMRI, skin conductance, and behavioral data during processing of concealed information. *Human Brain Mapping*, 28, 1287–301.

Genovese, C.R., Lazar, N.A., Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage*, 15, 870–8.

Gödert, H.W., Rill, H.G., Vossel, G. (2001). Psychophysiological differentiation of deception: the effects of electrodermal lability and mode of responding on skin conductance and heart rate. *International Journal of Psychophysiology*, 40, 61–75.

Greely, H.T., Illes, J. (2007). Neuroscience-based lie detection: the urgent need for regulation. *American Journal of Law & Medicine*, 33, 377–431.

Hakun, J.G., Seelig, D., Ruparel, K., et al. (2008). fMRI investigation of the cognitive structure of the Concealed Information Test. *Neurocase*, 14, 59–67.

Herron, J.E., Henson, R.N.A., Rugg, M.D. (2004). Probability effects on the neural correlates of retrieval success: an fMRI study. *Neuroimage*, 21, 302–10.

Horneman, C.J., O'Gorman, J.G. (1985). Detectability in the card test as a function of the subject's verbal response. *Psychophysiology*, 22, 330–3.

Iidaka, T., Matsumoto, A., Nogawa, J., Yamamoto, Y., Sadato, N. (2006). Frontoparietal network involved in successful retrieval from episodic memory. Spatial and temporal analyses using fMRI and ERP. *Cerebral Cortex*, 16, 1349–60.

Janisse, M.P., Bradley, M.T. (1980). Deception, information and the pupillary response. *Perceptual and Motor Skills*, 50, 748–50.

Kiehl, K.A., Laurens, K.R., Duty, T.L., Forster, B.B., Liddle, P.F. (2001). Neural sources involved in auditory target detection and novelty processing: an event-related fMRI study. *Psychophysiology*, 38, 133–42.

Konishi, S., Wheeler, M.E., Donaldson, D.I., Buckner, R.L. (2000). Neural correlates of episodic retrieval success. *Neuroimage*, 12, 276–86.

Kozel, F.A., Revell, L.J., Lorberbaum, J.P., et al. (2004a). A pilot study of functional magnetic resonance imaging brain correlates of deception in healthy young men. *Journal of Neuropsychiatry and Clinical Neurosciences*, 16, 295–305.

Kozel, F.A., Padgett, T.M., George, M.S. (2004b). A replication study of the neural correlates of deception. *Behavioral Neuroscience*, 118, 852–6.

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*, 605–13.

Langleben, D.D., Schroeder, L., Maldjian, J.A., et al. (2002). Brain activity during simulated deception: an event-related functional magnetic resonance study. *Neuroimage*, *15*, 727–32.

Langleben, D.D., Loughead, J.W., Bilker, W.B., et al. (2005). Telling truth from lie in individual subjects with fast event-related fMRI. *Human Brain Mapping*, *26*, 262–72.

Lee, T.M.C., Liu, H., Tan, L., et al. (2002). Lie detection by functional magnetic resonance imaging. *Human Brain Mapping*, *15*, 157–64.

Lee, T.M.C., Liu, H., Chan, C.C.H., Ng, Y., Fox, P.T., Gao, J. (2005). Neural correlates of feigned memory impairment. *Neuroimage*, *28*, 305–13.

Linden, D.E., Prvulovic, D., Formisano, E., et al. (1999). The functional neuroanatomy of target detection: an fMRI study of visual and auditory oddball tasks. *Cerebral Cortex*, *9*, 815–23.

Lykken, D.T. (1959). The GSR in the detection of guilt. *Journal of Applied Psychology*, *43*, 385–8.

Lykken, D.T. (1998). *A Tremor in the Blood: Uses and Abuses of the Lie Detector,* 2nd edn. New York: Plenum Press.

Nakayama, M. (2002). Practical use of the concealed information test for criminal investigation in Japan. In: Kleiner, M., editor. *Handbook of Polygraph Testing*. San Diego: Academic Press, pp. 49–86.

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

Nichols, T., Brett, M., Andersson, J., Wager, T., Poline, J. (2005). Valid conjunction inference with the minimum statistic. *Neuroimage*, *25*, 653–60.

Nuñez, J.M., Casey, B.J., Egner, T., Hare, T., Hirsch, J. (2005). Intentional false responding shares neural substrates with response conflict and cognitive control. *Neuroimage*, *25*, 267–77.

Nyberg, L., Marklund, P., Persson, J., et al. (2003). Common prefrontal activations during working memory, episodic memory, and semantic memory. *Neuropsychologia*, *41*, 371–7.

Phan, K.L., Magalhaes, A., Ziemlewicz, T.J., Fitzgerald, D.A., Green, C., Smith, W. (2005). Neural correlates of telling lies: a functional magnetic resonance imaging study at 4 Tesla. *Academic Radiology*, *12*, 164–72.

Raine, A., Lencz, T., Bihrle, S., LaCasse, L., Colletti, P. (2000). Reduced prefrontal gray matter volume and reduced autonomic activity in antisocial personality disorder. *Archives of General Psychiatry*, *57*, 119–27.

Ranganath, C., Johnson, M.K., D'Esposito, M. (2003). Prefrontal activity associated with working memory and episodic long-term memory. *Neuropsychologia*, *41*, 378–89.

Rosenfeld, J.P., Cantwell, B., Nasman, V.T., Wojdac, V., Ivanov, S., Mazzeri, L. (1988). A modified, event-related potential-based guilty knowledge test. *International Journal of Neuroscience*, *42*, 157–61.

Rosenfeld, J.P. (2002). Event-related potentials in the detection of deception, malingering, and false memories. In: Kleiner, M., editor. *Handbook of polygraph testing*. San Diego: Academic Press, pp. 265–86.

Rosenfeld, J.P., Soskins, M., Bosh, G., Ryan, A. (2004). Simple, effective countermeasures to P300-based tests of detection of concealed information. *Psychophysiology*, *41*, 205–19.

Rosenfeld, J.P., Biroschak, J.R., Furedy, J.J. (2006). P300-based detection of concealed autobiographical versus incidentally acquired information in target and non-target paradigms. *International Journal of Psychophysiology*, *60*, 251–9.

Russ, M.O., Mack, W., Grama, C., Lanfermann, H., Knopf, M. (2003). Enactment effect in memory: evidence concerning the function of the supramarginal gyrus. *Experimental Brain Research*, *149*, 497–504.

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

Spence, S.A., Farrow, T.F., Herford, A.E., Wilkinson, I.D., Zheng, Y., Woodruff, P.W. (2001). Behavioural and functional anatomical correlates of deception in humans. *Neuroreport*, *12*, 2849–53.

Strange, B.A., Henson, R.N., Friston, K.J., Dolan, R.J. (2000). Brain mechanisms for detecting perceptual, semantic, and emotional deviance. *Neuroimage*, *12*, 425–33.

van Hooff, J.C., Brunia, C.H., Allen, J.J. (1996). Event-related potentials as indirect measures of recognition memory. *International Journal of Psychophysiology*, *21*, 15–31.

Vasey, M.W., Thayer, J.F. (1987). The continuing problem of false positives in repeated measures ANOVA in psychophysiology: a multivariate solution. *Psychophysiology*, *24*, 479–86.

Venables, P.H., Christie, M.J. (1980). Electrodermal activity. In: Martin, I., Venables, P.H., editors. *Techniques in Psychophysiology*. Chichester: Wiley, pp. 3–67.

Verschuere, B., Crombez, G., De Clercq, A., Koster, E.H.W. (2004). Autonomic and behavioral responding to concealed information: differentiating orienting and defensive responses. *Psychophysiology*, *41*, 461–6.

Wagner, A.D., Desmond, J.E., Glover, G.H., Gabrieli, J.D. (1998). Prefrontal cortex and recognition memory. Functional-MRI evidence for context-dependent retrieval processes. *Brain*, *121*, 1985–2002.

Wolpe, P.R., Foster, K.R., Langleben, D.D. (2005). Emerging neurotechnologies for lie-detection: Promises and perils. *The American Journal of Bioethics*, *5*, 39–49.