A BEHAVIOURAL AND COGNITIVE NEUROSCIENCE INVESTIGATION OF DECEPTION COMMUNICATION

by

CHUN-WEI HSU

A thesis submitted to Plymouth University
in partial fulfilment for the degree of

DOCTOR OF PHILOSOPHY

School of Psychology

2018

Copyright Statement

Acknowledgements

I would like to express my deepest gratitude to my principal supervisor, Dr. Giorgio Ganis, for his constant support, encouragement and expert advice. His guidance helped me overcome numerous obstacles I have been facing throughout my research. It was a pleasure to work with him on a topic that was of great interest to me.

I would like to thank the rest of my supervisory team: Dr. Haline Schendan, Dr. Michaela Gummerum, and Prof. Clara Casco, for their insightful comments and feedback. I would also like to thank Prof. Sue Denham, the director of the CogNovo programme, for her great effort at creating so many unique opportunities of personal and professional development for CogNovians. I would also thank Sarah and Jacob for the help of data collection. Moreover, I would like to thank Art and Josh for their valuable advice from time to time.

I am extremely grateful to know all my Cognovian fellows. It has been a privilege to work with them, support each other and have fun time together. A special thanks goes to Michael for his playful thinking, stimulating me in many different ways, including a swear word in a talk. I would like to express my gratitude to Frank for his technical support. I am also indebted to Mihaela, Diego, and Raluca for their friendship and encouragement during the journey.

I would also like to extend my deepest gratitude to my family and friends from my home country for their remote support although most of them have no idea about my work. A special thanks goes to Hsiao Ma whose mentally and physically support have given me the strength to keep positive through my Ph.D. years. Moreover, I would like to thank my best friends, Liang Chun and A Lan, who have kept my spirits up.

AUTHOR'S DECLARATION

At no time during the registration for the degree of Doctor of Philosophy has the author been registered for any other University award without prior agreement of the Graduate Sub-Committee.

The study in Chapter 5 was conducted in collaboration with Professor Clara Casco during the author's 4 months secondment at University of Padova, Italy.

Work submitted for this research degree at the Plymouth University has not formed part of any other degree either at Plymouth University or at another establishment.

This study was financed with the aid of a fellowship from the People Programme (Marie Curie Actions) of the European Union's CogNovo (*FP7-PEOPLE-2013-ITN-604764*).

Presentation and Posters:

Talk on "the effect of mental countermeasures on fMRI-based concealed information tests" at 2nd Conference of Decepticon, Stanford University, USA, August 2017

Poster on "an ERP index of concealed information" at 23rd Annual meeting of Cognitive Neuroscience Society, New York, USA, April 2016

Poster on "creative cognition and deceptive communications" at 1st Conference of Decepticon, University of Cambridge, UK, August 2015

Talk on "creative cognition and deceptive communications" at 7th Annual School of Psychology Conference, Plymouth University, June 2015

Poster on "neurofunctional evidence for the dual process theory on value-based decisions" at 22nd Annual meeting of Cognitive Neuroscience Society, San Francisco, USA, March 2015

Word count for the main body of this thesis: 47,727

Signed ________________________

Date ________________________

A Behavioural and Cognitive Neuroscience Investigation of Deceptive Communication

Chun-Wei Hsu

ABSTRACT

There is a rich literature on how people tell lies and detect them in others, but the underlying mechanisms are still poorly understood. The first aim of this thesis was to elucidate key cognitive and neural processes underlying cued (i.e., instructed) and uncued lies. The second aim, based on recent research suggesting a link between dishonesty and creativity, was to determine whether creative cognition contributes to deceptive communication.

In a first behavioural study, performance on generating and detecting lies was measured in a socially interactive setting involving cued and uncued lies. Results of a multiple regression analysis showed that creativity predicted lying generation ability: more creative individuals were better liars than less creative people. In contrast, the ability to detect lies showed no association with creativity measures, suggesting that generating and detecting lies are distinct abilities.

A second event-related potential (ERP) study investigated the neural mechanisms underlying the generation of uncued lies using a novel bluffing paradigm where participants lied at will. Results showed no stimulus-locked differences between uncued lies and truths, suggesting that decision processes leading to both required comparable cognitive resources. Once the uncued decision has been made, it requires strategic monitoring to keep track of the responses in order to maximize the gains regardless of whether the outcome is a lie or the truth as indexed by no response-locked differences between uncued lies and truths.

Finally, parallel functional magnetic resonance imaging (fMRI) and ERP studies were conducted to determine the role of creativity in countermeasure use in a concealed information paradigm requiring cued lying. Results showed that countermeasures degraded the neural signatures of deception and more so for more creative individuals.

This work advances understanding of the cognitive and neural mechanisms underlying deception as well as their dependence on individual differences in creative cognition.

LIST OF CONTENTS

ACKNOWLEDGEMENTS V

AUTHOR'S DECLARATION VII

ABSTRACT IX

LIST OF TABLES XIII

LIST OF FIGURES XV

LIST OF ADDPENDICES XVIII

CHAPTER

I GENERAL INTRODUCTION

Deception…………………………………………………………………...1

Deception and Cognitive Processes……………………………………...2

Neurocognitive approach………………………………………………..…4

Deception cannot be Measured Directly……………………………......7

Paradigms Used to Examine Deception…………………………………9

Paradigms to investigate the neurocognitive processes underlying deception................…………………………. …………………………9

Instructed/cued lies……………………………………………………..10

Spontaneous/uncued lies……………………………………………….17

Paradigms to Detect Guilty Knowledge – Concealed Information Test…….22

CIT paradigms in ERP studies…………………………………………23

CIT paradigms in fMRI studies………………………………………...25

Countermeasures in the CIT paradigms………………………………..29

Deception and Creativity……………………………………………….31

Summary………………………………………………………………..36

II STUDY 1 – What Role Does Creative Cognition Play in Deceptive Communication?

Introduction……………………………………………………………….38

Materials and methods……..……………………………………………...44

Results…………………………………………………………………….54

Discussion………………………………………………………………….61

III STUDY 2 – An ERP Study of Uncued Lies in a Bluffing Game

Introduction……………………………………………………………………74
Materials and methods…..…………………………………………………...86
Results……………………………………………………………………...103
Discussion…………………………………………………………………..121

IV STUDY 3 – The Effect of Mental Countermeasures on ERP-based Concealed Information Test

Introduction………………………………………………………………...147
Materials and methods…………………………………………..…….…155
Results……………………………………………………………………...163
Discussion…………………………………………………………………..174

V STUDY 4 – The Effect of Mental Countermeasures on fMRI-based Concealed Information Test

Introduction………………………………………………………………...180
Materials and methods………..…………………………………………184
Results……………………………………………………………………...200
Discussion…………………………………………………………………..215

VI GENERAL DISCUSSION

Summary of findings………………………………………………………232
The Difference in P300 between the Bluffing and CIT Paradigms…………..233
Countermeasure Effects with ERP and fMRI Methods………………………234
Limitations and Future Work………………………………………………237
Conclusions………………………………………………………………...239

APPENDICES……………………………………………………………………..241

LIST OF ABBREVIATIONS………………………………………………………266

REFERENCES……………………………………………………………………268

LIST OF TABLES

CHAPTER 2 (STUDY 1)

Table 2-1 Descriptive statistics……………………………………………………..57

Table 2-2 Summary correlations and results from the stepwise regression analysis results……………………………………………………………………58

CHAPTER 3 (STUDY 2)

Table 3-1 Mean (SD) response time ...……………………………………………...105

Table 3-2 Paired t-tests of response time (RT)……………………………………106

Table 3-3 Correlations between d'_{Sender}, $d'_{Receiver}$, creativity scores (i.e., CAQ and ATTA) and the number of Total Lie……………………………………...119

CHAPTER 4 (STUDY 3)

Table 4-1 Mean (SD) response time (ms) in the NK, CK and CM conditions………165

Table 4-2 Lateral (Lat) and Midline (Mid) ANOVAs on N2 (200, 250 ms) and P300 (500, 700 ms) amplitude in no knowledge (NK), concealed knowledge (CK), countermeasure (CM) tasks, comparison between the CK and CM tasks and comparison between the CK and NK tasks. The item factor includes two levels: the probe and the average of the 4 irrelevants……………………173

CHAPTER 5 (STUDY 4)

Table 5-1 ANOVAs comparing response time (RTs) and error rates (ERs) for the probe and the irrelevants in all pairs of conditions. For the NK condition, the probe was only nominally defined because participants had no concealed information…………………………………………………………..202

Table 5-2 Peak MNI coordinates with Brodmann's Area (BA) and minimal t-statistic of brain regions with significant positive and negative responses in the contrast of Probe minus Irrelevants in the NK, CK and CM conditions. Note that for the NK condition, the probe was only nominally defined because participants had no concealed information………………………………..206

Table 5-3 Peak MNI coordinates and minimal t-statistic of brain regions showing significantly larger probe-irrelevants between conditions (CK/NK, CM/NK, and CK/CM) …………………………………………………………...207

Table 5-4 ANOVAs comparing activation for the probe and the irrelevants across all 7

ROIs and in all pairs of conditions (CK/NK, CM/NK, and CK/CM). Only significant effects involving the factors of interest Condition and Item are shown. Note that for the NK condition, the probe was only nominally defined because participants had no concealed information……………...210
Table 5-5 Area under the curve (AUC) for the comparisons CK vs. NK and CM vs. NK for 7 ROIs……………………………………………………………213

LIST OF FIGURES

CHAPTER 2 (STUDY 1)

Figure 2-1 Parameters for Sender and Receiver based on Signal Detection Theory (SDT) ………………………………………………………………..53

Figure 2-2 Sender's performance was predicted by the number of lies told when given a choice (Choice-Lie) and by CAQ-performance using this equation: *Predicted* $d'_{sender} = (-0.036 \times \text{Choice-Lie}) + (-0.011 \times \text{CAQ-performance}) + 0.616$ ($R^2 = .127$) ……………………………….58

Figure 2-3 Receiver's performance was not predicted by the regression model that successfully predicted Sender's performance using this equation: *Predicted* $d'_{receiver} = (-0.036 \times \text{Choice-Lie}) + (-0.011 \times \text{CAQ-performance}) + 0.616$ ($R^2 = .003$)……………………..59

Figure 2-4 Correlation between Sender and Receiver's performance using SDT measures ($r = -.099$, $p = .345$) …………………………………………60

Figure 2-5 Correlation between Sender and Receiver's performance using SDT measures in Opinion Survey (OS) task without Choice cue ($r = .003$, $p = .973$) ……………………………………………………61

CHAPTER 3 (STUDY 2)

Figure 3-1 Card stimuli used in the bluffing game………………………………...87

Figure 3-2 List of payoff possibilities in the bluffing game………………………91

Figure 3-3 Diagrams showing the sites analysed for the MFN, PRP and LPC….. 102

Figure 3-4 (a) Raster diagrams (Stimulus-locked) illustrating significant differences between ERPs to Cued truth and Small lie and between ERPs to Cued truth and Uncued truth. Note that there was no significant difference between ERPs to Uncued truth and Small lie so it is not shown here. Grand-average ERPs for the Cued truth, Uncued truth and Small lie condition at Fz with the N200 time window shown in grey (b) and at Pz with the P300 time window shown in grey (c)……………………..108

Figure 3-5 (a) Raster diagrams (Response-locked) illustrating significant differences between ERPs to Cued truth and Small lie and between ERPs to Cued truth and Uncued truth. Note that there was no significant difference between ERPs to Uncued truth and Small lie so it is not shown here.

Grand-average ERPs for the Cued truth, Uncued truth and Small lie condition at Cz (c) and Pz (d). Another Grand-average ERPs at Fz and Cz with BP time windown shown in grey (b)……..……………………..112

Figure 3-6 (a) Raster diagrams (Stimulus-locked) illustrating significant differences between ERPs to Lie detection and Head movement tasks at visual component (i.e., mean amplitude of Oz, O1, O2, Pz, PO3, PO4, P7 and P8). Topographic maps showing the difference between the Lie detection and Head movement tasks between (c) 100 and 300 ms and (d) between 400 and 700 ms (clip onset, collapsing stimuli)………………………117

Figure 3-7 Scatter plot between d'_{Sender} and $d'_{Receiver}$……………………………120

Figure 3-8 Scatter plot between the amount of Small lies and lie acceptability score…………………………………………………121

Figure 3-9 A diagram depicting some of the suggested cognitive processes for lying compared with truth telling……………………………………...…146

CHAPTER 4 (STUDY 3)

Figure 4-1 Digit stimuli…………………………………………………………...156

Figure 4-2 Grand-average ERPs for the probe and irrelevants at Pz as well as the scalp distribution maps (500, 700 ms) for the P300 probe effect in the (a) NK, (b) CK and (c) CM conditions…………………………………...166

Figure 4-3 Grand-average ERPs for the difference between the probe and irrelevants at Pz as well as the P300 scalp distribution maps (500, 700 ms)……..169

Figure 4-4 Mean potential (µV) for the probe and irrelevants in the CK and CM conditions at Pz…………………………………………………………..169

Figure 4-5 Grand-average ERPs for the probe and the irrelevants at Fz as well as the N2 scalp distribution maps (250, 350 ms) in the (a) NK, (b) CK and (c) CM conditions……………………………………………………...171

CHAPTER 5 (STUDY 4)

Figure 5-1 (a) Reaction times and (b) error rate for the probe and the irrelevants in the NK, CK and CM conditions………………………………………203

Figure 5-2 Whole brain map in the contrast of Probe minus Irrelevants in the CK (red) and CM (green) conditions (Note that for the NK condition, no

significant region was found) ……………………………………...205
Figure 5-3 Brain activation for the probe and the irrelevants across all 7 ROIs in the CK, CM and NK condition……………………………………………210

LIST OF APPENDICES

APPENDIX A Ethical approval of research……………………………………..241

APPENDIX 2A Subject information and consent form for Study 1……………...244

APPENDIX 2B Opinion survey and episodic memory items used in Study 1…...246

APPENDIX 3A Subject information and consent form for Study 2……………...247

APPENDIX 3B Follow-up questions for Study 2………………………………...249

APPENDIX 3C Supplementary results for Study 2……………………….……...250

APPENDIX 5A Subject information and consent form for Study 4……………...264

CHAPTER 1

GENERAL INTRODUCTION

Deception

Deception is a social behaviour defined as a deliberate attempt to convince someone of something the liar believes is untrue (Vrij & Ganis, 2014). DePaulo, Kashy, Kirkendol, Wyer, and Epstein (1996) have shown that deception is a common behaviour, with people reporting that they tell lies in one out of every four social interactions.

Although deception is frequent in daily life, the majority of lies are social lies (DePaulo et al., 1996) told in order to maintain a good relationship with others, to avoid awkward situations or discussions, or to avoid damaging a person's confidence and self-esteem. For example, you would say, "The steak was awesome" instead of "The steak was very tough" to your mother although she overcooked it. However, sometimes lies can cause serious negative consequences and we would like to detect them. For example, in a crime event, the police detective wants to know whether the suspect's alibi is true or not. At a country's border, customs officers want to know whether travelers are using fake identity information. During a job interview, an employer wants to know whether the candidate is as capable as he/she claims. Detecting these sorts of lies successfully would benefit not only individuals but also society. Therefore, researchers have been investigating how people tell lies and how to detect lies in others from both theoretical and applied angles. The theoretical angle focuses on general mechanisms and processes underlying deception, e.g., theory of mind, working memory, inhibition, mental management, etc. In contrast, the applied angle aims at developing methods to detect lies and at examining the validity of these methods. It is important to keep an eye on both perspectives because applications must rely on theories and theories may be further improved by taking into account insights generated by applications.

This introduction will review cognitive mechanisms and neural processes underlying deceptive communication and how this knowledge is applied to deception paradigms. This thesis aims at touching upon both theoretical and applied ends of the deception research spectrum by investigating cognitive and neural processes underlying deception using a variety of behavioral and neuroimaging methodologies, and examining the validity of lie detection methods.

Deception and Cognitive Processes

Deception is a task that usually requires effortful executive processes, such as inhibition, working memory, and so on (Gombos, 2006). Successful deceit requires multiple cognitive processes. For example, a liar must take the other person's thoughts into consideration (theory of mind) to construct a new scenario (the lie) while maintaining actual information (the truth) in mind. Furthermore, the liar also has to ensure self-consistency by keeping track of follow-up conversations based on lies that have already been told. Developmental studies have provided evidence that several cognitive functions are necessary to deliver a false message successfully. For example, *Theory of mind* is a key social ability for an individual to be able to deceive (e.g., Wimmer & Perner, 1983) and 3-year-old children have difficulty deceiving compared with 4- or 5-year-old children (e.g., Peskin, 1992) because theory of mind abilities emerge around age 4. In addition to the acquisition of specific abilities for lying, deceptive ability in children is closely related to the maturity of executive functioning, especially inhibition (Carlson, Moses, & Hix, 1998; Hala & Russell, 2001). For example, a study has shown that inhibiting a truthful response is an essential ability for children to be able to deceive (Carlson et al., 1998). In sum, research on deception and on the development of deceptive abilities suggests that social and executive control processes play a central role in performing deception.

Neurocognitive approach

In the past decade, researchers have begun to explore the neurocognitive basis of deception with the aim of understanding the connection between deceptive mental states and brain activity. Johnson (2014) reviewed a few dozen neurocognitive studies of deception and proposed three main categories of cognitive processes underlying deception: working memory, long-term memory, and executive/cognitive control processes.

Working memory, a central executive function with limited capacity, is responsible for maintaining and manipulating information, which are important for decision-making and for selecting goal-directed behaviors. It plays a vital role in carrying out deception, including integrating past memory with current information and fabricating self-consistent deceptive scenarios. The dorsolateral prefrontal cortex (DLPFC) is the most relevant brain areas for working memory and this region has also been found to be involved in deception (Abe et al., 2008; Abe et al., 2006; Bhatt et al., 2009; Ganis, Kosslyn, Stose, Thompson, & Yurgelun-Todd, 2003; Ito et al., 2012; Ito et al., 2011; Lee et al., 2009; Nunez, Casey, Egner, Hare, & Hirsch, 2005; Spence et al., 2001).

Most lies are about past events or knowledge stored in long-term memory.

Episodic and semantic memories are the most relevant to generate lies because they are consciously accessible sources about one's past and general knowledge. For example, in a crime situation, a suspect may lie about crime events and about memories related to the crime. Most neurocognitive studies of deception have used episodic, autobiographical or personal semantic memories as concealed information or as the stimuli for lying. For example, Ganis et al., (2003) obtained detailed information about participants' actual experiences in the past week as stimuli and found that the lies elicited additional activity in a number of brain areas involving in episodic memory retrieval, e.g., bilateral anterior prefrontal cortex, parahippocampal gyrus, and right precuneus.

Executive and cognitive control processes are also key for generating deceptive responses. Inhibiting a prepotent truthful response and making a deceptive response require executive functions to select and control actions, resolve response conflicts, and generate a response. During deception, a variety of executive processes that rely on different neural circuits act together. For example, the anterior cingulate cortex (ACC) plays an essential role in response monitoring and error detection. Several event-related potential (ERP) studies have found that instructed lies elicit a significantly larger medial frontal negativity (MFN), an ERP component peaking about 70 ms after a response

(Johnson, Barnhardt, & Zhu, 2004, 2005). The neural generators of the MFN have been localized to the ACC and adjacent cortex (Gehring & Willoughby, 2002; Johnson et al., 2004; Ullsperger & Von Cramon, 2001), which has been linked to response conflict monitoring (Gehring & Willoughby, 2002; Johnson et al., 2004; Mathalon et al., 2003; Vidal, Hasbroucq, Grapperon, & Bonnet, 2000).

Vrij and Ganis (2014) proposed a similar classification, with three classes of neural processes engaged during deception based on fMRI studies: cognitive control, memory-related, and social cognitive processes. As already mentioned, cognitive control processes are required to sustain the higher complexity of generating lies, including working memory, conflict monitoring, and response inhibition. Memory-related processes reflect the fact that lies and truths are ultimately based on stored memories. Finally, successful deception requires one to keep track of the perspective of the other person. As mentioned previously, a broad network of brain regions relevant to executive and memory-related functions is found in deception studies, including large portions of the prefrontal cortex, ACC and inferior parietal lobule (IPL). Recently, more interactive experimental paradigms have been used where participants had to take into account the perspective of a person and make a self-determined decision to deceive the other. Lisofsky, Kazzer, Heekeren, and Prehn

(2014) performed a quantitative meta-analysis in which the studies were divided into two categories, i.e., social-interactive and non-interactive deception studies. Social-interactive studies showed stronger activation in the right temporo-parietal junction (TPJ)/angular gyrus and the bilateral temporal pole (TP) than non-interactive deception. These results indicated that in addition to executive processes, perspective taking, theory of mind and moral reasoning processes also play an important role in deception. This meta-analysis also pointed out that social cognitive aspects should be taken into account when developing experimental paradigms for deception research.

The first two studies in this thesis included this socio-cognitive element in the experimental tasks, so as to be able to study deception situations closer to real life. By increasing the ecological validity of the paradigms, we were able to investigate some of the cognitive and neural processes underlying deception in socially interactive settings.

Deception cannot be measured directly

Deception originates from mental states, which are not directly observable. Thus, lie detection relies on an indirect route, that is, on a set of measureable variables that are correlated with deceptive mental states. For example, as deception usually requires greater cognitive resources than truth telling, much literature has shown that some behavioral cues associated with cognitive effort correlate with generating lies. The first meta-analysis of behavioural deceptive cues, Zuckerman, DePaulo, and Rosenthal (1981) indicated that deceptive speech is associated with more speech hesitations, pupil dilations and fewer hand movements. Ekman (2009) reported that facial expressions are associated with deceit. These facial displays (micro expressions) are associated with emotional or behavioral inhibition due to being deceptive. Vrij and Mann (2001) examined videotaped interviews with a convicted murderer and the murderer's behavior showed more gaze aversion, longer pauses, and slower and more frequent speech disturbances while lying compared with truth telling. These behavioral cues were attributed to greater cognitive load during deceptive responses.

Several variables have been investigated for lie detection in the literature. Vrij and Ganis (2014) summarized four general approaches to detect lies, which are non-verbal behaviors, verbal behaviors, physiological responses, and brain /neural activity. Unlike Pinocchio's growing nose, no unique response has been found to be robustly associated with deception. Such lack of robust associations is due in part to the fact that i) non-deceptive mental states may also correlate with these variables, resulting in false-positives (i.e., innocent individuals incorrectly classified as deceptive), and ii) deceptive mental states may correlate with other variables that we are not measuring, resulting in false-negatives (i.e., deceptive individuals incorrectly classified as innocent). These are the reasons why detecting deception is difficult and why the accuracy of the lie-detecting methods is not perfect and requires more examinations. This introduction will focus on the neural signatures of deception measured by electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) as these two methods are used in Studies 2, 3 and 4.

Paradigms used to examine deception

In the last two decades, researchers have begun to study directly the brain, where deception ultimately begins. The goals have been both to establish neurocognitive theories of deceptive processes and to develop better methods for detecting deception using a variety of paradigms. The paradigms can be broadly classified into two categories. The first category aims at revealing the brain-behavior relationships underlying deception. The second category aims to detect the presence of guilty knowledge in a person's memory. The following sections will review these two categories separately.

Paradigms to investigate the neurocognitive processes underlying deception

To understand the cognitive and neural correlates of deception and to develop general theories of deception, a number of neuroimaging studies employed paradigms to elicit cognitive processes associated with deceptive responses (e.g., Ganis & Keenan, 2009; Johnson, Barnhardt, & Zhu, 2003; Johnson et al., 2004, 2005; Lisofsky et al., 2014; for reviews: Sip, Roepstorff, McGregor, & Frith, 2008). A paradigm type referred to as “instructed lies” (it is also called “cued lies” in the current thesis) was used in many earlier neuroimaging studies where participants were instructed to lie about specific statements according to a cue. In this type of experimental setting, lying requires simply reversing responses, e.g., responding “no” for “yes”. Therefore, there has been much discussion about the validity of this type of paradigms because lies are not instructed in a real life. As the main feature of deception is to deliberately induce a false belief, more sophisticated types of deception paradigm involving spontaneous lies (it is also called “uncued lies” in the current thesis) have been considered. The following two sections will introduce paradigms using cued and uncued lies, respectively.

Instructed/cued lies

Recently, a few studies have investigated deception using paradigms grounded in the cognitive approach (Suchotzki, Crombez, Smulders, Meijer, & Verschuere, 2015). The fundamental idea is that lying is cognitively more demanding than telling the truth

(DePaulo et al., 2003; Vrij, Fisher, Mann, & Leal, 2006). Generating a deceptive response is a complex process, as it requires: (a) retrieving information associated with truthful and deceptive responses from memory, (b) maintaining both truthful and deceptive information in working memory, (c) inhibiting prepotent truthful responses, and (d) monitoring and controlling one's behavior to avoid revealing deceptive intentions (Ganis & Keenan, 2009; Vrij et al., 2006).

To elicit deceptive responses, a number of fMRI studies instructed participants to answer incorrectly in one condition and to tell the truth in another. For example, in the Sheffield Lie test developed by Spence et al. (2001), participants were presented with autobiographical events and instructed to lie or tell the truth depending on a cue with simple yes/no responses. Nunez et al. (2005) instructed participants to lie about autobiographical (e.g., "Do you own a laptop computer?") and non-autobiographical events (e.g., "Is a laptop computer portable?"), and compared the differences between these two types of lies. Kozel et al. (2005) created a mock-crime scenario where they instructed participants to "steal" a watch or ring and place it among their personal belongings in a locker. Subsequently, participants were asked questions about whether they took the item or not, which they had to deny.

Despite the variety of deception paradigms, Christ, Van Essen, Watson, Brubaker, and McDermott (2009) identified brain regions consistently showing deception-related activity across studies by utilizing an activation likelihood estimate (ALE) meta-analysis. These regions included bilateral ventrolateral prefrontal cortex (VLPFC), bilateral dorsolateral prefrontal cortex (DLPFC), bilateral anterior insula, anterior cingulate cortex (ACC) and bilateral inferior parietal lobule (IPL). To examine which cognitive processes are involved deception, they also compared the results of the deception meta-analysis with those of separate meta-analyses for three main executive control processes: working memory (e.g., keeping truth active while formulating a deceptive response), inhibitory control (e.g., inhibiting a truthful response), and task switching (e.g., switching between truthful and deceptive responses on different questions). The results showed that working memory plays a critical role in deception as deception-related regions in left DLPFC, right anterior PFC and right posterior parietal cortex were uniquely associated with working memory compared with inhibitory control and task switching. This type of work provides evidence about some of the neurocognitive basis of deception.

ERPs have also been used to study the cognitive processes underlying deception as they emphasize the temporal characteristics of the cognitive processes involved in deception. For example, Fang, Liu, and Shen (2003) instructed participants to deceptively deny having knowledge of familiar target faces, and to truthfully admit having knowledge of other familiar faces. In the study by Dong, Wu, and Lu (2010), participants were asked to evaluate the attractiveness of facial photos according to difference cues (i.e., truthfulness or deceptiveness). Johnson et al. (2004, 2005) instructed participants to make truthful and deceptive responses about old (remembered) and new (perceived) stimuli where participants had to sometimes correctly and sometimes incorrectly indicate recognition of old stimuli. In a study by Johnson, Henkell, Simon, and Zhu (2008), participants made truthful and directed lie (i.e., press opposite of the truth) responses about their attitudes towards moral issues (either strongly agreed or disagreed).

Consistent with the idea that deception requires additional executive processes, even simple instructed lies (reversing responses) produced a pattern of ERP activity that differed from that of truthful responses. There are specific ERP components that previous studies have discussed. For example, the Contingent Negative-going Variation (CNV; Brunia, van Boxtel, & Böcker, 2012; Walter, Cooper, Aldridge, McCallum, & Winter, 1964) is a slow negative brain potential occurring after a cue and before a stimulus which is regarded as indexing a process of anticipation and response preparation. An enhanced CNV has been reported for lying compared to truth telling

(Dong et al., 2010; Fang et al., 2003; Suchotzki, et al., 2015) and this CNV deception effect has been interpreted as an indication of increased cognitive load for lying compared to truth telling. The fronto-central N200 is another negative-going component occurring between 200 and 350 ms after stimulus onset. It has been considered to reflect the engagement of executive control processes, especially those related to conflict detection (Folstein & Van Petten, 2008; Van Veen & Carter, 2002). An increased fronto-central N200 has been found for lying compared to truth telling, as lying comes with increased response conflict and conflict inhibition (Hu, Wu, & Fu, 2011; Suchotzki et al., 2015; Wu, Hu, & Fu, 2009). A positive-going component, the P300, occurring between 300 and 800 ms after stimulus onset is found mostly over parietal scalp sites (Sutton, Braren, Zubin, & John, 1965). The P300 has been mostly studied in concealed information oddball paradigms where a larger P300 usually reflects the increased engagement of attention and memory processes for infrequent novel or salient stimuli (Polich, 2012). The P300 has also been considered as an indication of cognitive load (Isreal, Chesney, Wickens, & Donchin, 1980; Kramer, Wickens, & Donchin, 1985; Wickens, Kramer, Vanasse, & Donchin, 1983) and its amplitude is inversely related to the level of cognitive load. Consistent with this idea, previous deception studies have found a decreased P300 for lying compared to truth telling, as lying enhances cognitive load (Hu et al., 2011; Suchotzki et al., 2015; Wu et al., 2009).

A family of response-locked ERP components has also attracted attention in deception research, as deception requires continuously monitoring actions for conflicting response tendencies and errors (i.e., tactical monitoring) (e.g., Botvinick, Nystrom, Fissell, Carter, & Cohen, 1999; Carter et al., 1998), and sometimes it requires the use of long-term (i.e., strategic) monitoring processes to ensure that the selected responses meet one's plans and goals (Johnson et al., 2003, 2004, 2008). For example, the Medial Frontal Negativity (MFN; also referred to as Correct Response Negativity, CRN) is a fronto-central negative component that occurs between within 100 ms after an incorrect response (Falkenstein, Hohnsbein, Hoormann, & Blanke, 1991, 1995; Gehring, Goss, Coles, Meyer, & Donchin, 1993). Some studies have shown that the MFN is generated in or near the ACC (Gehring & Willoughby, 2002; Johnson et al., 2004; Ullsperger & Von Cramon, 2001), a brain region involved in response monitoring and conflict detection (e.g., Bush, Luu, & Posner, 2000; Johnson et al., 2004; Turken & Swick, 1999; Ullsperger & Von Cramon, 2001; Van Veen & Carter, 2002). Deceptive responses have been found to elicit a larger MFN than truthful responses, which has been attributed to stronger strategic/response monitoring demands for lying (Dong et al., 2010; Johnson et al., 2004, 2005, 2008; Kireev, Pakhomov, & Medvedev, 2008).

Another component, the pre-response positivity (PRP), unfolding between 250 and 350 prior to the response, has been found to be reduced for self-generated lies compared to directed lies, as self-generated lies require more strategic monitoring to be successful (Johnson et al., 2008). The parietal Late Positive Component (LPC), between 400 ms prior to and 100 ms after a response, provides a sensitive index of how processing resources are allocated between two simultaneously performed tasks (Johnson, 1993, 1998). In deception studies, a smaller LPC has been found for deceptive than for truthful responses (Johnson et al., 2003, 2005, 2008), because processing resources are allocated to secondary task (i.e., generating a lie) while holding the information of primary task (i.e., the truth). In sum, these ERP components index a variety of executive processes (e.g., conflict detection, response monitoring) and their modulation by deception manipulation provides evidence that deception is associated with increased cognitive load to control and coordinate the selection and execution of a non-default action.

These fMRI and ERP studies have shown the neurocognitive basis of deception using instructed/cued lie paradigms, focusing on executive control processes. However, these deception paradigms did not necessarily elicit the same processes that more ecological paradigms would as in real-world settings deception is a deliberate act that is intended to induce a false belief and mislead another person.

Spontaneous/uncued lies

With the increasing popularity of neuroimaging techniques to investigate deception, there has also been increased discussion about the validity of the deception paradigms employed in this field. Instructing participants to lie in one condition and to tell the truth in another condition does not fully satisfy the definition of deception — intentionally misleading another person. To simulate real life deception in the laboratory, Abe, Suzuki, Mori, Itoh, and Fujii (2007) were one of the first groups that tried to address the issue of “instructed lies” by introducing a clever twist to their experimental protocol. In their paradigm, participants received two contradictory instructions from two experimenters where experimenter 2 secretly asked the participants to deceive the experimenter 1 by answering questions with opposite responses to those told the experimenter 1 when experimenter 1 was not present. However, the issue of “instructed lies” was still unsolved because participants were just instructed to follow a different set of rules.

To make the experimental paradigms more ecologically valid, some authors attempted to include “uncued lies” in their studies. For example, Spence, Kaylor-Hughes, Farrow, and Wilkinson (2008) requested participants to write down their true events (embarrassing autobiographical episodes) that a typical person would probably have wished to concealed from others and participants were free to choose when to tell the truth or to lie. The results showed that VLPFC was associated with lying compared with truth telling. This paradigm captured real-world deception by granting participants choices to lie or tell the truth about these embarrassing events. However, there are other aspects of deception that are important as deception is also socially rooted and its processing is modified by moral perception (Sip et al., 2008). In other words, to execute a deceptive act depends on the existence or expectation of interactive consequences and no harm would come without such consequences.

Recently, a few fMRI studies have recognized that deception involves a social dimension and considered the pragmatics of social interaction into their experimental designs. Baumgartner, Fischbacher, Feierabend, Lutz, and Fehr (2009) used a trust game where participants made a promise whether they were going to give back money to an interactive partner and then kept or broke the promise. Their findings revealed that breaking a promise was associated with increased activation in the DLPFC, ACC and amygdala. The authors suggested that the dishonest act involves an emotional conflict due to the suppression of the honest response. Sip et al. (2010) developed a laboratory version of the dice game where participants decided whether or when to attempt to deceive their opponent, and whether or when to accuse them of deception. The results showed that both truthful and deceptive responses were associated with activity in fronto-polar cortex (BA 10) while deceptive responses were associated with greater activity in the premotor and parietal cortex. In another study by Sip et al. (2012) the paradigm created a context in which deceptive responses could be confronted by the opponent because in real-life the cost of being caught can be enormous in terms of loss of money, reputation, or trust. Their findings suggest that deception depends upon an effort-based affective-motivational network (e.g., subgenual ACC) rather than merely higher-level cognitive processes and the decision to deceive is affected by the potential risk of social confrontation rather than the claim itself.

More ecologically valid and interactive experimental paradigms seem to yield different neural networks from what previous studies have found using instructed lies. To investigate the influence of these socio-cognitive elements on the neural networks underlying deception, Lisofsky et al. (2014) performed a quantitative meta-analysis combining the data from 416 participants across 22 fMRI and two PET studies. These studies were divided into social interactive and non-interactive deception categories based on the description of the experimental paradigm. The results showed that social interactive deception was associated with greater activation in the dorsal ACC, the right temporo-parietal junction (TPJ)/angular gyrus, and the bilateral temporal pole (TP) than non-interaction deception. Activity in dorsal ACC has been found in dilemmatic moral reasoning (Greene, Nystrom, Engell, Darley, & Cohen, 2004). The TPJ has been associated with theory of mind processes where people integrate socially significant information and infer the mental state of others (e.g., Saxe & Kanwisher, 2003; Saxe, Moran, Scholz, & Gabrieli, 2006) while the TP has been consistently found to be associated with moral judgment and social cognition (for a review, see Olson, Plotzker, & Ezzyat, 2007). Thus, instead of a neural network involved in executive control processes (for a review, see Christ et al., 2009), this meta-analysis demonstrated the critical role of theory of mind, perspective taking, moral reasoning and conflict processes in socially interactive deception studies.

A few ERP studies have also developed tasks to characterize the electrophysiological signature of spontaneous/uncued deception. Instead of the ERP components mentioned in the previous section, another component related to moral conflict was found in a more ecologically valid paradigm. Panasiti et al. (2014) implemented a two-card (i.e., winning card and losing card) game, in which an opponent (OP) had to pick one of the two covered cards and the participants had to communicate the outcome of the choice to the OP because the OP was not able to see the outcome of the choice. Participants were completely free to decide whether to lie or not to the OP and the winner took money from the other player. This paradigm captured two important components of deception, willful decision and social interaction. Instead of supporting the consensus that deception is cognitively more demanding than truth telling, this study suggested that spontaneous deception, as opposed to instructed lies, might not mandate additional cognitive workload by showing neither P300 nor N200 differences between lying and telling the truth. Interestingly, the study found a reduction in the Bereitschaftspotential (BP) for lying compared with truth telling. The BP is a volitional movement marker modulated by low-level variables, such as motor parameters, or by high-level cognitive variable, such as decision-making based on the mental computation of the trade-off between morality and reward. Panasiti et al. (2014) attributed this reduced BP to moral conflict as lying took the reward from the opponent. This paradigm close to real life successfully elicited moral emotions (e.g., guilty and shame) in deception demonstrating that it is important to use ecologically valid paradigms to approach realistic deception. However, there was still one component of deception missing in this paradigm — there was no risk or punishment if participants lied. In other words, participants would never be caught when they made deceptive responses. This thesis developed a card game paradigm in Study 2 that involved the elements of deception mentioned above to examine the electrophysiological signature of spontaneous/uncued deception.

Paradigms to detect guilty knowledge – Concealed Information Test

These paradigms were developed during the second half of the 20th century using physiological measures for detecting deception, primarily in crime cases. Lykken (1959) initially developed the Guilty Knowledge Test (GKT) based on the idea that a perpetrator possesses specific crime information that an innocent person does not. Both guilty and innocent people provide the same behavioral response by denying having such information, but the perpetrator lies when doing so. Thus, it is a form of deception related to cued lies for those individuals with guilty knowledge.

Recently, the GKT has been referred to as the Concealed Information Test (CIT; see Verschuere, Ben-Shakhar, & Meijer, 2011). The CIT utilizes a series of multiple-choice questions, in which a critical item (often referred to as the probe) is presented together with some neutral/control items (often referred to as irrelevant items). These control items are chosen to be of a similar type as the probe so that an innocent person would not be able to discriminate them from the probe (Lykken, 1998). Thus, the relevant item (i.e., probe) is significant only for guilty individuals. In addition, there is a target item where examinees have to press a "Yes" button. The reason to have the target item is to ensure that the examinees do not simply press a “No” button mindlessly throughout the study.

For example, in a murder case, the suspects could be asked, “Was the murder weapon a knife? Was the murder weapon a gun? Was the murder weapon a rope?” Although a guilty person denies having such knowledge by answering no to all these questions, orienting responses (ORs) still occur. The OR is a set of physiological and behavioral reactions that reflect attentional processes (e.g., Dawson, Filion, & Schell, 1989; Filion, Dawson, Schell, & Hazlett, 1991; Siddle & Spinks, 1992), and are elicited by novel stimulus or by any change of stimulation (e.g., Berlyne, 1960; Sokolov, 1963). Traditionally, the ORs are indexed by autonomic nervous system (ANS) measures, such as increased skin conductance response (SCR), respiration suppression and heart rate deceleration (for reviews, Ben-Shakhar & Elaad, 2003; Meijer, Selle, Elber, & Ben-Shakhar, 2014). In more recent years, neuroimaging measures have frequently been employed to examine concealed information (e.g., Farwell & Donchin, 1991; Ganis, Rosenfeld, Meixner, Kievit, & Schendan, 2011; Rosenfeld et al., 1988). The following two sections will focus on ERP and fMRI, respectively.

CIT paradigms in ERP studies

Event-related potential (ERP) studies using the 3-stimulus CIT paradigm just described have used mostly the amplitude of the P300 component as a robust index of recognition of concealed knowledge, with the probe showing a larger P300 than the irrelevant items (e.g., Farwell & Donchin, 1991; Ganis, Bridges, Hsu, & Schendan, 2016; see Meijer et al., 2014 for reviews; Rosenfeld, Soskins, Bosh, & Ryan, 2004). The P300, a positive brain potential that occurs between 300 and 800 ms after stimulus onset was first described by Sutton et al. (1965). Typically, the P300 component is found in oddball paradigms, for example when detecting a rare stimulus (e.g., high pitch tone) interspersed in a series of standard stimuli (e.g., low pitch tones).

This neural index is based on the fact that the amplitude of the P300 is inversely related to the perceived category probabilities of a stimulus (see Johnson, 1988 for reviews). In the CIT paradigm, the relevant guilty knowledge item (i.e., probe) is random and infrequent (normally 16.7%) compared with other irrelevant items (66.7%). The idea is that, if a person possesses the guilty knowledge, the infrequent probe item will be highly salient, and it will automatically stand out from the rest of the irrelevant items and elicit a larger P300. Conversely, if a person has no knowledge of the guilty information, the probe will be treated like the irrelevant items and elicit the similar amplitude of P300 as irrelevant items do. Although some habituation effects on the P300 have been found (Polich, 1989), the P300 does not habituate as fast as the skin conductance response (SCR) (Elton, Schandry, & Sparrer, 1983; Rushby & Barry, 2009). This allows for repeating each stimulus many times, resulting in a better signal-to-noise ratio than autonomic nervous system (ANS) measures (Meijer et al., 2014).

With regard to the accuracy of the ERP method, Meijer et al. (2014) conducted a meta-analysis on the validity of the CIT and found that the P300 has the largest effect size (Cohen's *d,* 1.89) compared with other three measures, i.e., skin conductance response (1.55), respiration line length (1.11) and heart rate (0.89). This review study concluded that P300 is a highly valid measure for detecting concealed information.

CIT paradigms in fMRI studies

Compared with ERP studies, a relatively small number of fMRI studies have been conducted so far using the CIT paradigm and the designs varied in different studies. Langleben et al. (2002) and Phan et al. (2005) used the same card stimuli but found slightly different patterns of results. Langleben et al. (2002) found stronger activations in the medial prefrontal cortex (medPFC), anterior cingulate cortex (ACC) and sensorimotor areas in the left hemisphere including the inferior parietal lobule (IPL), but no activation was found in the ventrolateral prefrontal cortex (VLPFC). Using the same stimuli but unlike the previous study, Phan et al. (2005) counterbalanced the probe and irrelevant cards across participants. Stronger activation was found in the dorsal medial prefrontal cortex (not including the ACC), VLPFC, right superior temporal sulcus and IPL for the probe compared with irrelevant cards. Both studies interpreted these activated brain regions as reflecting increased engagement of response monitoring and inhibition processes during deception. Another two fMRI studies (Gamer, Bauermann, Stoeter, & Vossel, 2007; Gamer, Klimecki, Bauermann, Stoeter, & Vossel, 2009) used the CIT protocol with cards and bank notes as stimuli. No counterbalancing of stimuli across participants was applied. They found more activation in the right insula and adjacent inferior frontal cortex (IFG) and the right middle cingulate gyrus for probe than irrelevant items. No differences were found in the ACC. However, these results can be confounded by differential intrinsic properties of probe and irrelevant items, since there was no counterbalancing of stimuli between participants. Furthermore, no control group with uninformed participants was used in these studies. In this thesis, proper counterbalancing of stimuli and a control condition with uninformed participants were used (see Study 3 and 4).

In another study, Nose, Murai, and Taira (2009) used a CIT paradigm based on playing cards and included a control group with no concealed knowledge. Compared with irrelevant items, the probe showed stronger activation in the bilateral VLPFC, left

IFG, right middle frontal gyrus and right IPL, but no difference was found in the ACC. Also, no differences were found between probe and irrelevant items in the control group. The right VLPFC, the most robust difference between with and without concealed information, was attributed to the need to inhibit recognition that may reveal knowledge of the probe. Another CIT study used dates as stimuli (Ganis et al., 2011) where participants were instructed to lie about their date of birth (probe). A control condition (no-knowledge) was included, within-subject where only irrelevant and target dates were presented. Numerous areas showed stronger activation for probe than irrelevant items in the concealed knowledge condition, including bilateral VLPFC, bilateral medPFC, ACC, middle cingulate gyrus and bilateral IPL. The medPFC and VLPFC showed the most robust differences, which was interpreted as memory retrieval and novelty detection processes.

The fundamental mechanism underlying the CIT to detect concealed knowledge is based on the orienting response (OR). However, fMRI studies have shown that another mechanism might contribute to the CIT effect, which is response inhibition, as several studies have found more neural activity in the rIFG (e.g., Gamer et al., 2007; Gamer et al., 2009; Ganis et al., 2011; Nose et al., 2009; Phan et al., 2005). The rIFG is considered an area crucially involved in inhibitory processes as it has been consistently found in traditional response inhibition tasks (Aron, Robbins, & Poldrack, 2004, 2014; Derrfuss, Brass, Neumann, & Von Cramon, 2005; Kelly et al., 2004). To directly examine the effect of response inhibition in the CIT, Suchotzki, Verschuere, Peth, Crombez, and Gamer (2015) instructed participants to admit knowledge of half of the mock crime-related probe items and deny the other half and compared the neural activities between deny and admit condition. They found that activation in the rIFG was significantly larger when deception was involved in (i.e., deny condition). Thus, it indicated that response inhibition is a crucial mechanism that drives the blood oxygen level dependent activity (BOLD) responding to concealed information. In addition, Suchotzki et al. (2015) also manipulated the proportion of probe versus irrelevant items (1:1 vs. 1:4) and found that the BOLD CIT effect was more robust in the 1:4 condition compared with 1:1 condition. To attain the largest effect, the 1:4 proportion of probe versus irrelevant items was used in this thesis (Study 3 & 4).

The accuracy of fMRI-based methods is an important issue as companies have begun marketing fMRI-based “lie detection” services. Davatzikos et al. (2005) used high-dimensional non-linear pattern classification methods (support vector machines [SVMs]) to discriminate between patterns of entire brain associated with lie and truth. 99% of the true and false responses were discriminated correctly and predictive accuracy assessed by cross-validation (leave 1% out from training a classifier) showed 88% accuracy. Monteleone et al. (2009) conducted one-out single-participant analyses and found that the medPFC had the best classification between deceptive and honest cases, which could identify 71% of participants as lying with no false alarms. Nose et al. (2009) performed a one-out cross-validation analysis based on the activation of the right VLPFC and they could classify individuals with and without concealed information with 84% accuracy. Ganis et al. (2011) used a one-out approach and linear SVMs based on the three regions (i.e., the left and right VLPFC and medPFC) and showed 100% accuracy to discriminate concealed knowledge and non-concealed knowledge cases. Langleben et al. (2016) directly compared the accuracy between fMRI and polygraph experts in a within-subjects study. Using a logistic regression, they found that fMRI experts were 24% more likely to detect the concealed information than the polygraph experts.

Countermeasures in the CIT paradigms

Despite the relative accuracy of neuroscience-based methods in detecting concealed knowledge, poor specificity is a big issue when applying these methods to the real world. For the ERP-based method, the P300 (the key index used to detect concealed knowledge) is elicited by any type salient and infrequent stimuli. In other words, P300 enhancement is not only produced by concealed information in the CITs, and so usually it can be susceptible to simple countermeasures. For example, Rosenfeld et al. (2004) used covert physical acts (e.g., imperceptibly pressing the left forefinger on the leg to an irrelevant item) as countermeasures and hit rates dropped from 82% to 18%. Additional ERP research on countermeasure will be discussed further in Study 3.

For the fMRI-based method, a meta-analysis encompassing all CIT studies to date (Ganis et al., in preparation) found stronger activation for concealed knowledge (i.e., the difference between probe and irrelevant items) in a network of regions similar to that found in the deception meta-analyses discussed earlier, including the bilateral IFG/insular cortex, the ACC, and bilateral IPL. These activated regions are not specific for concealing information, as they overlapped with those found in meta-analyses of executive processes: similar brain activation patterns can be generated by cognitive processes that have nothing to do with concealing information. This poor specificity makes the fMRI methods potentially vulnerable to countermeasures as well. Indeed, a previous fMRI study with the CIT using the participant's date of birth as the probe found reliable effects of countermeasures (Ganis et al., 2011). The countermeasure used in this study was training participants in associating specific covert actions with specific irrelevant items (e.g., imperceptibly wiggling the left index finger upon seeing the first irrelevant item). The accuracy for detecting deception was 100% without countermeasures, but the accuracy for classification dropped to 33% (4 out of 12) when using the countermeasures. The effect of countermeasure on fMRI methods to detect concealed knowledge will be discussed further in Study 4.

Given that these countermeasures can be implemented easily, additional research is needed to examine the vulnerability of neuroimaging methods to various countermeasures before these methods can start to be used to detect deception in the real world. This thesis addressed this question by conducting two studies to examine how a novel mental countermeasure affects the validity of neuroimaging methods to detect concealed information. The effect of mental countermeasures on P300 amplitude was examined in Study 3, and on fMRI activation in Study 4.

Deception and Creativity

As mentioned earlier, although most lies are not serious, deception sometimes can cause serious consequences and bring high costs for individuals and society. Thus, the topic of deception has attracted a lot of attention and investment among law enforcement, polygraphers and forensic psychologists. Although there is a rich literature on how people tell lies and detect them in others, the role of individual differences in deceptive abilities have not been sufficiently investigated. Following up on recent research suggesting a bidirectional link between dishonesty and creativity (Gino & Ariely, 2012; Gino & Wiltermuth, 2014), this thesis aimed at exploring the role of creative cognition in deceptive communication.

People are expected to tell the truth all the time; thus, being dishonest involves breaking this social rule. On the other hand, to be creative, one must break standard rules to take advantage of existing opportunities, or to create new ones (Brenkert, 2009). Given that both dishonesty and creativity encompass rule breaking, these two abilities may facilitate each other. Indeed, there are many examples of "malevolent creativity" from the history of warfare by implementing innovative tactics and technologies in order to achieve victory. A number of examples in organizational settings also showed "negative creativity" by finding novel ways of stealing from a company. These creative activities are beneficial to one group but not another. This concept can also be applied to terrorism where terrorists create surprising products that are effective in achieving their goals (Cropley, Kaufman, & Cropley, 2008). A novel act of terror can be regarded as creative, despite the fact that the results do not benefit our common good.

Recently, some empirical studies have provided evidence for this link between dishonesty and creativity. Gino and Wiltermuth (2014) conducted five experiments where participants had the opportunity to behave dishonestly by overreporting their performance on various tasks and then they completed several tasks designed to measure creativity. One of their experiments included two supposedly unrelated tasks: a computer-based math-and-logic game and the Remote Associate Test (RAT). In the first task, participants were randomly assigned to either the control or the likely-cheating condition. In the control condition, participants completed the task without further instructions. In the like-cheating condition, participants were told that the computer had a programming glitch where the correct answer would appear on the screen unless they stopped it from being displayed by pressing the space bar right after the problem appeared. The experimenter also informed participants that they should try to be honest and solve the problems on their own although no one would be able to tell whether they had pressed the space bar or not. In reality, the “glitch” was a feature of the program and the number of space-bar presses was recorded and used as a measure of cheating. After the math-and-logic game, participants completed 12 RAT problems, which measured creativity by assessing people’s ability to identify associations between words. For example, each problem consists of a set of three words (e.g., sore, shoulder, sweat), and participants have to find a word that is logically linked to them (cold). The results showed that most participants (51 out of 53) cheated in the likely-cheating condition and their RAT performance (*M* = 6.20, *SD* = 2.72) was higher than those in the control condition (*M* = 4.65, *SD* = 2.98), $t_{(97)} = 2.71$, $p = .008$. These results demonstrated that cheating increased creativity in a subsequent task.

In another study, Gino and Ariely (2012) used multiple measures of creativity as well as various tasks to provide participants with the opportunity to cheat. They found that a positive correlation between creative personality and the level of dishonesty. Participants' intelligence was also measured but no link was found between creativity and intelligence or between dishonesty and intelligence. Another experiment of Gino and Ariely (2012) examined whether activating a creative mindset temporarily can promote dishonest behavior by randomly assigning participants to either a creative-mindset or a control condition. All participants completed a scrambled sentence test where they had to construct grammatically correct four-word sentences (e.g., "The sky is blue") from a set of five randomly positioned words (e.g., sky, is, the, why, blue). In the creative-mindset condition, 60% of words related to creativity (e.g., creative, original, novel, new, etc.), but no words related to creativity in the control condition. Subsequently, participants completed one problem-solving task allowing them to overstate their performance. The results showed that people who were primed to think creatively were more likely to behave dishonestly than those in the control condition.

A study on malevolent creativity conducted by Beaussart, Andrews, and Kaufman (2013) showed a negative link between behavioral integrity and creativity. The result showed that more creative individuals would tend to receive the reward immediately without completing the task, rather than go back and complete the task even though they could have received the reward without doing so. In addition, one recent study has shown a robust positive relationship between broad foreign experiences and immoral behavior across a variety of cultural populations using multiple methods, including longitudinal, correlational, and experimental (Lu et al., 2017). They suggested that foreign experiences enhance cognitive flexibility and creativity that lead to immoral behavior by increasing moral relativism, i.e., the belief that morality is relative rather than absolute.

Thus, previous work has shown that breaking rules may be the root of both dishonest behavior and creative performance. However, the potential link between deceptive communication (i.e., the ability to generate lies and to detect lies) and creativity remains unknown. As creative thinking allows individuals to be flexible (Runco, 2004) and to adapt ideas according to the context in order to achieve an original and suitable solution (Brophy, 1998, 2001), creative cognition may facilitate generating lies by enabling them to flexibly build connections between previously unassociated elements. Thus, people with higher creativity may generate more plausible and convincing lies. On the other hand, creativity has been found to be related to "leaky attention", meaning that some information from non-attended sources leaks in and is processed (Zabelina, Saporta, & Beeman, 2015). Thus, creative cognition may facilitate detecting lies using behavioural cues as well by filtering out fewer deceptive cues, enabling people to detect lies in others more accurately. These hypotheses will be examined mainly in Study 1.

Summary

This introduction reviewed general cognitive processes underlying deception, including cognitive control and social cognitive processes, especially from the perspective of neuroscience. Moreover, the bidirectional link between dishonesty and creativity was discussed providing the rationale for an individual differences approach to investigate the potential overlap between processes involved in deceptive communication and creative cognition. This was done in Study 1 by using a socially interactive setting involving cued and uncued lies.

Paradigms used to examine deception were also reviewed and subdivided according to two categories. One category of paradigms is used to investigate and understand the cognitive and neural basis of deception employing instructed/cued and spontaneous/uncued lies. Study 2 examined the neural activities underlying uncued deception in an interactive setting.

The other category of paradigms is used is to detect the presence of guilty knowledge in a person's memory and focuses on the applied end. Studies 3 and 4 aimed at examining the vulnerability of concealed information paradigm, which requires cued lying, when applying mental countermeasures using ERP and fMRI methodologies respectively, and determining the role of creativity in countermeasure use (see these chapters for specific hypotheses).

CHAPTER 2

STUDY 1 – What Role Does Creative Cognition Play in Deceptive Communication?

Introduction

This behavioural study aimed at measuring the individual differences in the ability to generate lies and to detect lies in a socially interactive setting involving cued and uncued lies and at determining how creative cognition contributes to these abilities.

Creative thinking allows individuals to solve problems in novel ways (Mumford & Gustafson, 1988) and to be flexible in order to cope with changes in their daily lives (Runco, 2004). Researchers have explored many cognitive processes important for creative cognition. Divergent thinking has been identified as one main component underlying creative performance and it refers to the ability of individuals to develop original ideas and to find multiple solutions to a given problem (Sternberg, 1999). Convergent thinking has also been reported as one cognitive process in creativity (Brophy, 2001; Finke, Ward, & Smith, 1992; Ward, Smith, & Vaid, 1997) and it involves judging and adapting ideas in order to achieve an original and appropriate solution (Brophy, 1998, 2001). Creative thinking entails thinking outside the box

(Runco, 2010) and it requires people to break standard rules within a domain to build associations between previously unassociated cognitive elements (Guilford, 1950). Breaking the rules and unusual mental associations can lead to novel ideas (Sternberg, 1988), but in the social domain it can also lead to dishonesty: being dishonest requires breaking the social norm that one should be honest all the time.

Recent research has provided evidence for a bidirectional link between creative behaviour and dishonesty because they both share this rule-breaking feature. In a study by Gino and Ariely (2012), more creative individuals tended to cheat more in laboratory tasks than less creative individuals. In addition, dishonest behaviour could be increased by priming participants to think creatively. Conversely, Gino and Wiltermuth (2014) found that individuals who cheated in a task were subsequently more creative than non-cheaters, even after taking into account individual differences in their creative ability. They also confirmed this link by randomly assigning participants to either the likely-cheating or the control condition. In the likely-cheating condition, a fake programming glitch allowed participants to see the correct answer during a math-and-logic game unless they stopped it from being displayed by pressing the space bar. The results support the idea that acting dishonestly leads to greater creativity in subsequent tasks. However, the potential role of creative cognition on the ability to generate lies and to detect them in others remains unknown. The main aim of this study was to investigate for the first time the relationship between creative cognition and deceptive communication, that is, the generation and detection of lies.

Developing a plausible lie requires cognitive resources to flexibly retrieve and manipulate episodic and semantic memories within the ongoing social context (Sip et al., 2008). Creative cognition may facilitate generating lies by enabling liars to flexibly restructure knowledge in different ways depending on the changing situational demands. Thus, creative individuals may be able to construct more creative and plausible deceptive scenarios. Creative cognition may also facilitate detecting lies by means of behavioural cues. People's ability to detect lies by using behavioural cues is just slightly better than chance on average, with only 54% of truths and lies being correctly classified (Bond & DePaulo, 2006). One reason for this is that cues to deception are typically weak and unreliable (DePaulo et al., 2003), so people overlook them easily or rely on stereotypical beliefs about deceptive cues that typically lead to poor lie detection performance. Creative cognition could facilitate detecting lies by enabling people to filter out fewer cues, and to rely less on stereotypical beliefs (Zabelina, Saporta, & Beeman, 2016).

Wright, Berry, and Bird (2012) proposed the existence of a "deception-general ability" by showing a significant positive correlation between the ability to generate and detect lies across people: in their study, better liars were also better at detecting lies. Thus, the second aim of the current study was to determine whether creative cognition might contribute to both generating and detecting lies, possibly accounting for the reported link between the ability to generate and to detect lies. Finding common effects of creative cognition on deception generation and detection would provide further evidence that they are two facets of a single deception-general ability. On the other hand, finding that creative cognition affects only deception generation or deception detection would provide evidence against the existence of such a deception-general ability.

In order to measure individual differences in both generating and detecting lies under an interactive and real-life environment, we implemented a group-based competitive deception game based on Wright et al. (2012): the Deceptive Interaction Task (DeceIT) where controversial topics were used as stimuli (e.g., "Smoking should be banned in all public places") and participants were required to make both truthful and deceptive statements about each topic. In this paradigm, groups of 5 participants competed with each other in both generating lies ("Sender") and detecting lies ("Receiver") so that both deceptive abilities could be simultaneously evaluated within participants. High-value prizes for the best performers in generating and detecting lies were given out in order to enhance participants' motivation.

In order to measure a person's ability to generate and to detect lies comprehensively, instead of using a single category of stimuli (i.e., controversial topics), our interactive deception game required participants to use their deceptive skills in three different domains: personal opinions using questions about controversial topics, personal experiences using episodic memory questions, and emotions using emotional photos. Participants' overall deceptive abilities were measured using combined results across these domains. Moreover, to enable participants to take the initiative in engaging in the deception task rather than being passive in receiving the cues (i.e., either tell a lie or a truth according to a cue), on some trials participants were allowed to choose whether they lied or told the truth (we will refer to these spontaneous lies as *uncued lies*). This design also allowed us to assess the relationship between frequency of uncued lies and the ability to generate believable lies and to answer the question of whether people are inclined to tell more lies if they are good at lying.

We used two creativity measures, the Creativity Achievement Questionnaire (CAQ) and the Remote Associate Test (RAT). The CAQ is based on the sum of creative products generated by an individual throughout their lifetime. Creative achievement seems to be facilitated by a confluence of factors, including the capacity for divergent thinking, imagination, intelligence, confidence, nonconformity and the ability to provide practical support (Amabile, 1996; Eysenck, 1995; Ludwig, 1995). The other measure, the RAT, was developed as a mean to measure creativity without requiring knowledge specific to any fields (Mednick, 1962). In this test, sets of three words are provided and participants have to find the solution associated with the three words (e.g., the three problem words “Dew/Comb/bee” are associated with the solution “Honey”). The RAT assesses associative abilities related to creativity and problem-solving. For the two creativity measures in the current study, the CAQ is an index of divergent thinking as the CAQ score has been found to be positively correlated with all facets (i.e., diversity, fluency, originality and flexibility) of divergent thinking (Carson, Peterson, & Higgins, 2005). On the other hand, the RAT is an index of convergent thinking as several reference suggested recently (Arden, Chavez, Grazioplene, & Jung, 2010; Benedek, Könen, & Neubauer, 2012; Kaufman, Kaufman, & Lichtenberger, 2011; Nielsen, Pickett, & Simonton, 2008). Indeed, success on the RAT is determined by whether the problem solver can identify the single and correct solution that meets the definition of convergent thinking (i.e., discerning which ideas are most appropriate or of highest quality and come up with a single and correct solution (Brophy, 2001; Guilford, 1967).

We expected to see a positive relationship between different aspects of creative cognition and the abilities to generate and detect lies. In addition, better liars were expected to choose to tell more lies than worse liars when given the choice to lie or not, as people would feel more confident to use the skill if they were good at it. Finally, we expected to replicate the main finding by Wright et al. (2012), that is, a positive correlation between the ability to generate and detect lies, and that both abilities would be equally affected by creative cognition.

Materials and methods

Participants

98 healthy native English speakers (66 females, 32 males, Mean age = 20.6 years, range: 18-35) without any difficulties in reading (i.e., no dyslexia) were recruited from the student population at Plymouth University, UK, and were either paid 8 pounds/hour for participating or received course credit. Informed consent was obtained from all participants. The full subject information and consent form is presented in Appendix 2A. The study was approved by the Faculty of Health and Human Sciences Ethics Committee.

Materials

Once participants were recruited, they completed an “Opinion Survey (OS)” online questionnaire, which included ten controversial topics, such as “Should we stop doing medical experiment on animals?” Participants responded “agree” or “disagree” on a 6-point scale and wrote down the reasons for their choice. These controversial topics were used as stimuli in the Opinion Survey (OS) task. All OS topics are presented in Appendix 2B. Just prior to the interactive deception game, participants answered ten questions about recent episodes from the daily life, such as “What did you do last weekend?” on paper. The ten questions were used as stimuli in the Episodic Memory Survey (EMS) task. All EMS topics are presented in Appendix 2B. In the Emotional Photos Description (EPD) task, twelve photos were selected from the International Affective Picture System (Lang, Bradley, & Cuthbert, 2005). Half of them were chosen to induce pleasant emotions (valence: 7.40 (±0.48); arousal: 6.54 (±0.86)) and the other half to induce unpleasant emotions (valence: 1.96 (±0.35); arousal: 6.83 (±0.36)).

In a second session, after the interaction deception game, participants completed the Wechsler Abbreviated Scale of Intelligence (Wechsler, 1999), creativity tests, i.e., Remote Associate Test (RAT) (Bowden & Jung-Beeman, 2003), the Creativity Achievement Questionnaire (CAQ) (Carson et al., 2005), and two psychometric tests, the Toronto Alexithymia Scale (TAS) (Taylor, Bagby, & Parker, 1992), and the Interpersonal Reactivity Index (IRI) (Davis, 1983). The details of each test/questionnaire are described next.

Wechsler Abbreviated Scale of Intelligence (WASI): The WASI (Wechsler, 1999) consists of four subtests: Vocabulary, Block Design, Similarities and Matrix Reasoning. This test was administered in about 30 minutes and testing time varied between subjects according to their performance.

Remote Associate Test (RAT): Mednick (1962) developed the Remote Associate Test (RAT) as a mean to measure creativity without requiring knowledge specific to any field. In this test, participants have to find a unique solution that is related to all three words provided. The relationship type can vary, including synonym, formation of a word, or semantic association. Bowden and Jung-Beeman (2003) created 144 items with a uniform way to approach the solution – formation of a compound word/phrase. For example, the three problem words "Manners/Round/Tennis" are associated with the solution "Table". We selected 105 out of 144 items in Bowden and Jung-Beeman (2003) excluding 39 items with solving rate of 1% or less in 2 seconds according to the Appendix in Bowden and Jung-Beeman (2003). There were four solution time limits (2, 7, 15 and 30 seconds) in Bowden and Jung-Beeman (2003) study. They found a 0.93

correlation between the percentage of participants producing solutions at 2 and 7 seconds. In order to minimizing testing time, we used 2 seconds as the solution time limit in the current study.

Creativity Achievement Questionnaire (CAQ): The CAQ is based on the sum of creative products generated by an individual throughout her/his lifetime. This CAQ self-report checklist consists 10 different areas of talents with total 96 items. Each ten standard domains or art (visual arts, music, dance, creative writing, architectural design, humor, and theatre and film) and science (culinary arts, inventions and scientific inquiry) contains eight ranked questions weighted with a score from 0 to 7. Participants are asked to place a checkmark on the description fitting their achievement best.

Toronto Alexithymia Scale (TAS): This TAS-20 self-report scale with 20 items measures the ability to identify and describe emotions in the self. Alexithymia refers to a condition in which people have trouble identifying and describing emotions and tend to minimize emotional experience and focus on attention externally. Items are rated using a 5-point Likert scale and the sum of responses to all 20 items is the total alexithymia score. The TAS-20 uses cutoff scoring: a score equal to or less than 51 is classified as non-alexithymia, a score equal to or greater than 61 is classified as alexithymia, and scores between 52 and 60 are classified as possible alexithymia.

Interpersonal Reactivity Index (IRI): This 28-item scale measures empathy, defined as the reactions of one individual to the observed experiences of another (Davis, 1983). This scale is answered on a 5-point Likert scale ranging from “Does not describe me well” to “Describe me very well”. There are four subscales with 7 different items for each (taken directly from Davis, 1983): Perspective taking – the tendency to spontaneously adopt the psychological point of view of others; Fantasy – taps respondents’ tendencies to transpose themselves imaginatively into the feelings and actions of fictitious characters in books, movies, and plays; Empathic concern – assesses “other-oriented” feelings of sympathy and concern for unfortunate others; Personal distress – measures “self-oriented” feelings of personal anxiety and unease in tense interpersonal settings.

As the TAS and the IRI scales provide self- and other-focused measures of emotional intelligence (Parker, Taylor, & Bagby, 2001), we were able to examine whether emotional intelligence would contribute the ability to generate lies by manipulating emotions to construct more convincing lies and to detect lies by understanding emotions in others to make a correct judgement.

Procedure

In the first session, participants were recruited in groups of five for a "Communication Skills Study" (except for two groups of four participants due to logistical reasons), with the constraint that people within the same group should not be acquainted with each other before the experiment. The reason for this requirement was to ensure participants did not know each other's background. This check was carried out through emails or text messages. At the beginning of the recruitment, in order to prevent preparation of lies before coming to the experiment, participants were told they would take part in a "Communication Skills" study instead of a "Deception Study". Each participant was assigned a number from one to five, which would become his or her identifier during the game. Participants took turns in making true or false statements as Senders while the other four people acted as Receivers, and were instructed to judge whether the Sender was making true or false statements. Each group completed three tasks. All participants in a group served as Senders in one task before moving on to the next task. The order of tasks for each group was counterbalanced.

The three tasks (OS, EMS and EPD) were presented using E-Prime 2.0. On each trial, the Sender sat in front of the rest of the group (the four Receivers) and watched a computer screen. A cue (i.e., the word 'Truth', 'Lie' and 'Choice') was presented for one second before each question or photo. For the OS and EMS tasks, a question appeared on the screen word by word, with each word presented for 300 milliseconds (ms) to equalize reading time across participants. For the EPD task, a photo appeared in the middle of the screen for 1500 ms. The sender had to make a true or false statement according to the given cue (Truth or Lie). When presented with the ‘Choice’ cue, Senders could freely choose whether to tell the truth or a lie. After seeing a question or an emotional photo, Senders were instructed to speak for up to 20 seconds (in the OS and EMS task) or 10 seconds (in the EPD task) as soon as they were ready. Next, Receivers judged whether the Sender told the truth or lied. In total, across the three tasks (10 Truth, 10 Lie, and 12 Choice cues), each participant completed 32 trials (10 for OS, 10 for EMS and 12 for EPD) as Sender. Every question and photo was shown only once for each participant to avoid any potential effects of repetition.

Participants were instructed to appear as credible as possible as Senders, regardless of whether they told the truth or lied. In order to further engage participants, monetary incentives were offered to the best performers in the Sender and Receiver roles (£30 for the first place; £20 for the second place; £10 for the third place).

Before each task started, all Senders had 2 practice trials for the OS and EMS tasks, and 4 trials for the EPD task. In the first practice trial of the OS and EMS tasks, Senders were shown a verbatim example of a response to illustrate the type of statement required (OS: “I support immigration. Immigrants usually come from poor regions. They can find a better life in a more developed country. People in well-developed countries should have more empathy for the poor and give them a chance to have a better life”; EMS: “The best trip I have ever had was when I took a boat to a small island with my family. This small island was beautiful. We saw many landscapes and enjoyed a lot of fresh food cooked by the local residents. The weather was very sunny during those days”), and they generated their own statement in the second practice trial. In the EPD task, Senders were presented with 2 practice trials of pleasant photos and 2 of unpleasant photos. In the first pleasant and the first unpleasant photo, the Sender was presented with a verbatim example of a response and they generated their own statement in the second photo.

Participants came back to attend the second session individually for the WASI, the four questionnaires (TAS, IRI, and CAQ), and the computer-based RAT task.

Data collection and analysis

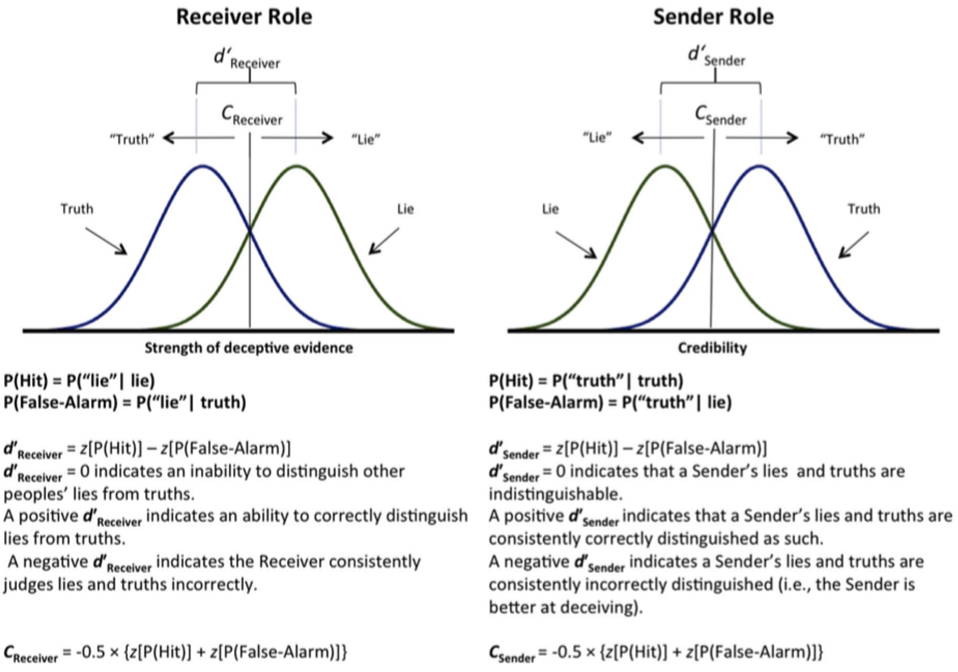
Performance during the Sender and Receiver roles were analyzed using Signal Detection Theory, SDT (Green & Swets, 1966). SDT allows lie-truth discriminability (*d’)* to be measured independently of judgment bias (*C*). SDT measures for Sender and

Receiver roles were calculated separately. These measures can be used when there are unequal proportions of lie and truth statements to be classified. For the Sender, a ”hit” (H) can be defined as a “truth” statement correctly judged as the truth, and a “false alarm” (FA) can be defined as a “lie” statement incorrectly judged as the truth. For the Receiver, a “hit” can be defined as a correct “lie” classification to a lie statement, and a “false alarm” can be defined as a incorrect “lie” classification to a truth statement. The discriminability of the Sender’s truths and lies is indexed by d'_{Sender} and the corresponding measure of bias, C_{Sender}, indicates the perceived credibility of a Sender regardless of the veracity. For the Receiver, d'_{Receiver} indicates the capacity of discriminating lies from truths whereas the corresponding measure of bias, C_{Receiver} indexes the tendency for a Receiver to endorse a given statement as truthful (credulity). The calculation of each parameter is described in the Figure 2-1 (From Figure 1, Wright et al., 2012). The SDT supposes that both lie and truth are normally (Gaussian) distributed with the same variance (i.e., the equal variance assumption). The discriminability (*d’)* is defined as the separation between the two means expressed in units of their common standard deviation.

With these indices, more successful deception as a Sender role is indicated by more negative values of d'_{Sender} and better lie detection as a Receiver role is indicated by more positive values of d'_{Receiver}. For the judgment bias (C), appearing more credible regardless of the veracity as a Sender role is indicated by more negative values of C_{Sender}, whereas showing more truth bias as a Receiver is indicated by more positive values of C_{Receiver}.

Deceptive abilities are quantified by d'_{Sender} and d'_{Receiver}, where d'_{Sender} represents the ability to generate credible lies whereas d'_{Receiver} represents the ability to detect lies in others.

Figure 2-1. Parameters for Sender and Receiver based on Signal Detection theory (SDT) (original resource: Figure 1, Wright et al. (2012))



P(Hit) = P("lie"| lie)
P(False-Alarm) = P("lie"| truth)

$d'_{\text{Receiver}} = z[\text{P(Hit)}] - z[\text{P(False-Alarm)}]$
$d'_{\text{Receiver}} = 0$ indicates an inability to distinguish other peoples' lies from truths.
A positive d'_{Receiver} indicates an ability to correctly distinguish lies from truths.
A negative d'_{Receiver} indicates the Receiver consistently judges lies and truths incorrectly.

$C_{\text{Receiver}} = -0.5 \times \{z[\text{P(Hit)}] + z[\text{P(False-Alarm)}]\}$

P(Hit) = P("truth"| truth)
P(False-Alarm) = P("truth"| lie)

$d'_{\text{Sender}} = z[\text{P(Hit)}] - z[\text{P(False-Alarm)}]$
$d'_{\text{Sender}} = 0$ indicates that a Sender's lies and truths are indistinguishable.
A positive d'_{Sender} indicates that a Sender's lies and truths are consistently correctly distinguished as such.
A negative d'_{Sender} indicates a Sender's lies and truths are consistently incorrectly distinguished (i.e., the Sender is better at deceiving).

$C_{\text{Sender}} = -0.5 \times \{z[\text{P(Hit)}] + z[\text{P(False-Alarm)}]\}$

Results

We collected data from 98 participants during the first session. Data from 97 of these participants was collected in the second session. During the analyses, it became clear that two participants did not understand the rules of the game (for example, they did not follow the cues to tell the truth or to lie in the interactive deception game), and so they were excluded. One participant was excluded because of a stammering problem. One outlier identified by more than three times the standard deviations in CAQ scores was also excluded. Therefore, the final analyses are based on 93 participants.

Deceptive abilities

Individual differences in deceptive abilities were examined by SDT analysis with four performance measures: $M\ d'_{\text{Sender}} = 0.463$, SD = 0.322; $M\ C_{\text{Sender}} = -0.329$, SD = 0.164; $M\ d'_{\text{Receiver}} = 0.441$, SD = 0.254; $M\ C_{\text{Receiver}} = 0.336$, SD = 0.191 (see Table 2-1). In the choice condition, participants told the truth much more frequently than they lied ($M = 9.0$, SD = 2.6 vs $M = 3.0$, SD = 2.6, respectively), a significant difference, $t(92) = 11.037$, $p < .001$, $d = 1.14$. Fourteen participants (15%) never lied in this condition.

Receiver performance was analyzed using conventional percentage accuracy rates, and overall accuracy was 60.2% (SD = 5.0%). This accuracy rate is significantly greater than chance ($t(_{92}) = 19.791$, $p < .001$, $d = 2.05$), but also significantly greater than the 54% reported in previous work ($t(_{92}) = 12.046$, $p < .001$, $d = 1.25$) (Bond & DePaulo, 2006). Fractional rates addressing accuracy for different types of statement showed a significantly higher mean accuracy for truths ($M = 70.7\%$, SD = 7.3%) than for lies ($M = 45.5\%$, SD = 9.3%, $t(_{92}) = 17.485$, $p < .001$, $d = 1.81$). The percentage of statements of all types classified by Receivers as truthful was 64.2% (SD = 6.6%), significantly greater than chance ($t(_{92}) = 20.586$, $p < .001$, $d = 2.13$).

Due to a technical problem, we were unable to record the response latency from participants of the last 10 groups. Response latency from the first 47 participants was 1252 ms (SD = 399) when they told the truth and 1306 (SD= 489) when they told the lies, but this difference was not significant ($t_{(45)} = -1.014$, $p = .316$).

Creativity measures and cognitive tests

Participants completed the Creativity Achievement Questionnaire (CAQ) based on the sum of creative achievements throughout one’s lifetime and the scores were found as follows (see Table 2-1): expressive domain (visual arts, writing and humor) of CAQ ($M = 4.86$, SD = 4.70); performance domain (dance, drama and music) of CAQ ($M =$

3.86, SD = 6.37); science domain (invention, science and culinary) of CAQ ($M = 2.27$, SD = 3.09). In the other creativity test, RAT, based on the ability of associating remote words quickly (2 seconds), participants answered 9.14 out of 105 items on average (range 0-29) (see Table 2-1).

The average score of Wechsler Abbreviated Scale of Intelligence (WASI) was 109.82 (SD = 9.86). The average score of Toronto Alexithymia Scale (TAS), a measure of the degree to which emotions can be identified and described by the individual, was 49.54 (SD = 10.61) with 16 (17.2%) participants classified as alexithymic and 22 (23.7%) participants classified as indeterminate. For the Interpersonal Reactivity Index (IRI), a measure of empathy, the average score was 69.89 (SD = 12.41).

Associations between deceptive abilities and creative cognition

A stepwise multiple linear regression was conducted to predict participants' ability to generate lies (d'_{Sender}) using 11 potential predictors (variables 2-12 in Table 2-1). The final model contained two of the eleven predictors and was statistically significant, $F_{(2, 90)} = 6.530$, $p < .005$, with an $R^2 = .127$ and an adjusted $R^2 = .107$ (see Table 2-2): Choice-Lie and CAQ-performance. The regression equation is shown in the legend of Figure 2-2. Both variables were negatively correlated with the ability to generate lies

(d'_{sender}), indicating that individuals with higher scores on these variables tended to have better ability to generate lies (the more negative the d'_{sender} and the better the Sender is at deceiving). The number of lies told when given a choice (Choice-Lie) had the largest weight, followed by CAQ-performance.

Table 2-1. Descriptive statistics

Variable	**Measure**	**Mean ± SD**
Deceptive abilities	1. d'_{sender}	0.463 ± 0.321
	2. C_{sender}	-0.329 ± 0.164
	3. d'_{receiver}	0.441 ± 0.254
	4. C_{receiver}	0.336 ± 0.191
	5. Choice-Lie[a]	3.04 ± 2.56
Creativity tests	6. CAQ-expressive[b]	4.86 ± 4.70
	7. CAQ-performance[c]	3.86 ± 6.37
	8. CAQ-scientific[d]	2.27 ± 3.09
	9. RAT[e]	9.14 ± 6.20
Intelligence	10.WASI[f]Full-IQ	109.82 ± 9.86
Psychometric tests	11.TAS[g]	49.54 ± 10.61
	12.IRI[h]	69.89 ±12.41

[a]The number of lies told when given the choice; [b]the Creativity Achievement Questionnaire, expressive domain score; [c]The Creativity Achievement Questionnaire, performance domain score; [d]The Creativity Achievement Questionnaire, scientific domain score; [e]Remote Associate test score; [f]WASI = Wechsler Abbreviated Scale of Intelligence; [g]Toronto Alexithymia Scale; [h]Interpersonal Reactivity Index; [i]MarlowCrowne Social Desirability scale

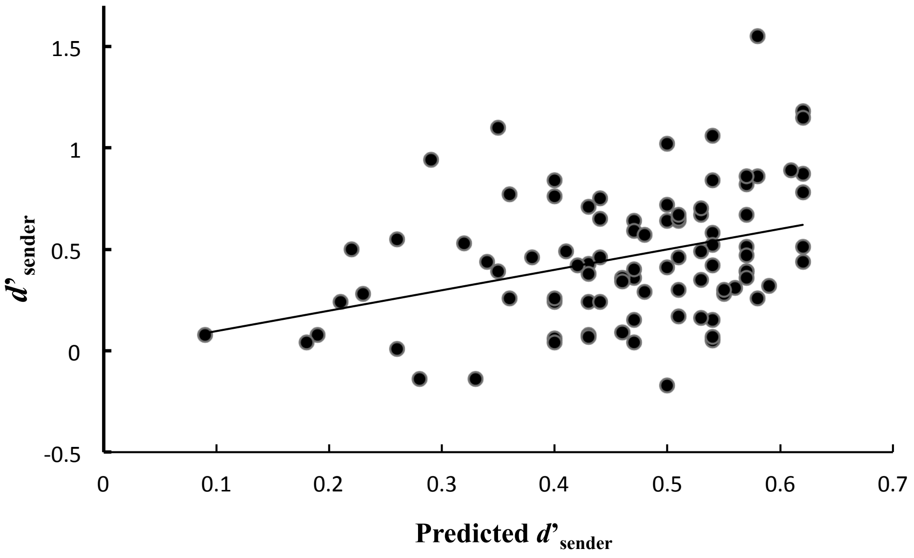
Table 2-2. Summary correlations and results from the stepwise regression analysis results

Variables	Correlation with d'_{Sender}	Multiple regression weights	
		b	β
d'_{Sender}			
Choice-Lie[a]	-.279**	-0.036**	-.287
CAQ-performance[c]	-.211*	-0.011*	-.221

Note. The dependent variable was d'_{Sender}, $R^2 = .127$ and Adjusted $R^2 = .107$;

* $p < .05$; ** $p < .01$

Figure 2-2. Sender's performance was predicted by the number of lies told when given a choice (Choice-Lie) and by CAQ-performance using this equation: *Predicted* $d'_{\text{sender}} = (-0.036 \times \text{Choice-Lie}) + (-0.011 \times \text{CAQ-performance}) + 0.616$ ($R^2 = .127$)

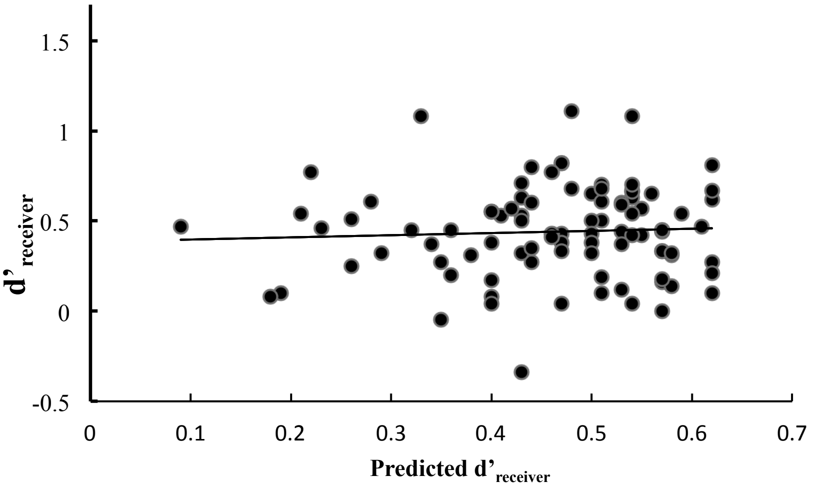


A second stepwise regression analysis was conducted to predict participants' ability to detect lies (d'_{receiver}) from the same11 potential predictors (except d'_{sender}

replaced *d*'$_{\text{receiver}}$ as a predictor). However, none of the variables survived the stepwise regression analysis. To confirm the specificity of the regression model found for generating lies, we used it also to try to predict lie detection ability. Results showed that the regression model for predicting lie generation ability did not predict lie detection ability (R^2 = .003, Figure 2-3).

In sum, higher lying ability is associated with telling more lies when given the choice to do so and with a higher performance score on the creativity achievement questionnaire, whereas the ability to detect lies is not associated with any of the potential predictors.

Figure 2-3. Receiver's performance was not predicted by the regression model that successfully predicted Sender's performance using this equation: *Predicted* d'$_{\text{receiver}}$ = (-0.036 x Choice-Lie) + (-0.011 x CAQ-performance) + 0.616 (R^2 = .003)



Next, we examined the relationship between the ability to generate lies (*d*'$_{\text{sender}}$) and the ability to detect lies (*d*'$_{\text{receiver}}$) in our dataset, as done in previous studies (Wright et al., 2012; Wright, Berry, Catmur, & Bird, 2015). There were no hints of a correlation between these two abilities ($r = -.099$, $p = .345$, see Figure 2-4). Moreover, in order to directly compare our findings with previous results, an additional correlation analysis between these two deception abilities was carried out using data from the Opinion Survey task only, excluding Choice cue trials. Again, there was no correlation between generating and detecting lies ($r = .003$, $p = .973$, see Figure 2-5).

Figure 2-4. Correlation between Sender and Receiver's performance using SDT measures ($r = -.099$, $p = .345$).

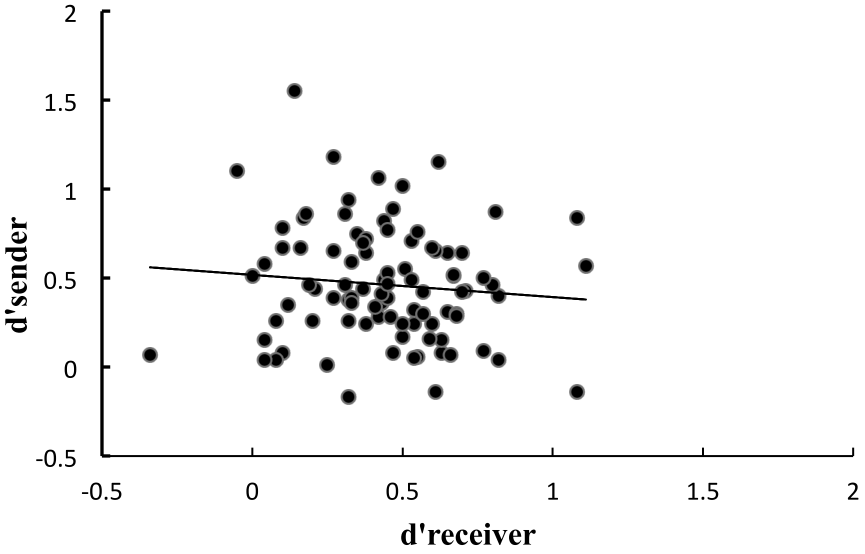
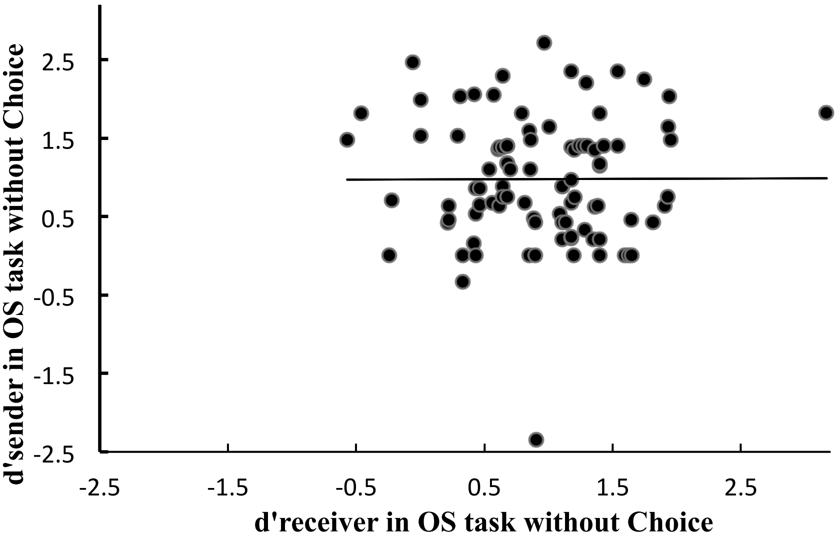


Figure 2-5. Correlation between Sender and Receiver’s performance using SDT measures in Opinion Survey (OS) task without Choice cue (*r* = .003, *p* = .973).



Discussion

Numerous studies have highlighted the importance of being creative for individuals, organizations and society. For example, investments in creativity and innovation can enhance organizational performance (Lev, 2004). Creative thinking helps us manage our daily life and generate novel ideas to solve problems. This study investigated another, possibly less positive side of creativity, by examining the contribution of different aspects of creative cognition to deceptive communication.

The finding of a positive association between the ability to generate lies and real-world creative achievement suggests that creative cognition shares some processes with deceptive communication. Creative cognition (especially divergent thinking), thinking outside the box and building links between previously unassociated elements, may facilitate generating credible lies. Furthermore, the ability to generate lies is significantly correlated with creative achievement in the performance domain (i.e. dance, drama and music) and not in the expressive or science domains. This indicates that creative individuals in the different domains may show different aspects of creative dispositions, and this has been supported by previous studies. For example, Feist (1998) conducted a meta-analysis and found that creative people in art and science exhibit different personality traits. Consistent with this, Silvia, Kaufman, and Pretz (2009) found three latent groups (i.e., a visual arts class, a performing arts class and a no creativity class) in the CAQ (n = 749) using latent class analysis. They suggested that the classes were meaningful by showing a significantly higher openness personality trait in the visual art and performing arts class, but a higher extraversion trait only in the performing arts class. Our findings also support this idea that creative individuals in the arts (especially in performance) and the sciences display different traits and different cognitive abilities. Here, we demonstrated that more creative individuals in the performance domain may be better at restructuring knowledge in multiple different ways to generate convincing lies with their higher imaginative ability. Also, being good at performance, such as dance, acting and music, may make it easier for a person to control their body language and hide deceptive cues.

The most significant variable to predict participants' ability to generate convincing lies in the regression model is the number of lies told when given the choice to do so (Choice-Lie). In other words, individuals who are good at generating plausible lies, choose to tell more lies when they are allowed to do so. One interpretation of this finding is that senders with a high *d*' may already know they are good at telling lies, and they may feel confident enough to choose to deceive others more often. Future studies can investigate whether these individuals also tell more lies in real life. If individuals with higher *d*' also show higher frequency of lying in their real life, it is possible that daily engagement in lying leads them to performing better also when lying in a laboratory study, and the rewards gained by successful lies encourage them to tell more lies when they have the chance to do so.

We hypothesized that creative cognition may also facilitate detecting lies by filtering out fewer potential deceptive cues that one could use to make the correct judgment. However, we failed to find any correlation between the ability of detecting lies and creative cognition. There are some possible reasons, described as follows. First, in the current interactive deception game, detecting lies required full attention throughout the whole experiment where participants had to make 32 judgments for each of the other four participants in the group. Thus, it is possible that the ability to detect lies might not be measured consistently due to difficulties in maintaining concentration throughout the whole game. Second, one meta-analysis has shown that people are barely able to detect lies using behavioral cues and the accuracy of a deception judgment depends more on the liar than on the judge (Bond Jr & DePaulo, 2008). Finally, personal learning and previous experience may play a more important role in detecting deception than creative cognition abilities as Ekman and O'sullivan (1991) found that one occupational group (i.e., Secret Service) was significantly more accurate than all of the other groups (e.g., federal polygraphers, judges, psychiatrists, college students, etc.).

The other creativity test, RAT, showed correlation with neither lying nor lie-detecting ability. The RAT (Mednick, 1962; Mednick & Mednick, 1967) has been used to examine the link between creativity and problem-solving, and it has been found to be related to analytical and deductive processing (e.g., intelligence, working memory and academic achievement) but only weakly correlated with divergent thinking scores (Lee, Huggins, & Therriault, 2014). In addition, there is a growing consensus that the RAT mainly measures convergent thinking processes (Brophy, 2001; Hommel, Todd, Hills, & Robbins, 2012). The current RAT performance results are consistent with the evidence reported in previous studies. First, the RAT was found to be significantly correlated with IQ scores ($r = .355$, $p < .001$), showing that it is more likely engaged in analytical processes related to intelligence. Second, the RAT performance showed no correlation with total CAQ score, the other creativity measure. This is consistent with the idea that RAT performance relates to convergent thinking processes that are different from the cognitive processes (e.g. divergent thinking and imagination) essential for creative achievement (Amabile, 1996; Eysenck, 1995; Ludwig, 1995); indeed, it has been proposed that divergent and convergent thinking are distinct cognitive abilities that represent unique facets of creative cognition (Lee & Therriault, 2013). On the other hand, creative achievement has been found to be correlated with divergent thinking test scores in one meta-analysis (Kim, 2008), and Plucker (1999) showed that divergent thinking strongly predicted creative achievements in a longitudinal study using structural equation modelling. The CAQ score is also substantially correlated with all facets (i.e., diversity, fluency, originality and flexibility) of divergent thinking (Carson et al., 2005). Thus, the ability to generate lies may be linked to divergent thinking processes captured in part by the CAQ but not by the RAT.

Unexpectedly, the two psychometric tests (i.e., TAS and IRI) showed contribution to neither lying nor lie-detecting ability. The TAS is used to examine whether people have trouble identifying and describing emotions while the IRI is used to examine perspective taking and empathy. No correlation between the TAS and deceptive abilities might indicate that the capacity for manipulating emotions was not involved in the deceptive abilities. On the other hand, although generating credible deceptive responses requires theory of mind (ToM) and perspective taking, the lack of a correlation with the IRI might suggest that in our paradigm even relatively low levels of ToM and perspective taking were sufficient for people to generate credible lies. A paradigm requiring more complex social interactions (e.g., involving several iterations), may be able to capture more subtle effects of ToM and perspective taking in generating lies.

The interactive deception game used in the current study was adapted from the Deceptive Interaction Task (DeceIT) by Wright et al. (2012). That study found a significant positive relationship between the ability to generate and detect lies, a finding that was replicated by Wright et al. (2015) in a later study. Based on this evidence, Wright et al. (2012) proposed the existence of a “deception-general ability” that would underlay both the ability to generate and detect lies. The authors of the study suggest that skills in both the generation and detection of deception would offer evolutionary advantages, and improvements in the ability to deceive in one species would produce selective pressure for corresponding improvements in deceptive detection among competitors and vice versa (Bond Jr & Robinson, 1988). Under this evolutionary selection idea, Wright et al. (2012) proposed that improvement in either lie generation or detection will result in concomitant improvements in the other ability. However, this argument may explain parallel changes in the ability to lie and to detect lies in a population over time but it is not clear how it could explain the correlation between these two abilities across participants *within a population*.

Indeed, we failed to find a significant positive correlation (or correlation at all) between these two deception abilities (r = -.099, p = .345, N=93) in the current study. Although the design and procedure was very close to that used in Wright et al. (2012) and had similar limitations, there were some differences that is important to discuss.

First, in our study there were three tasks, i.e., Opinions Survey, Episodic Memory Survey and Emotional Photo Description, compared with only one task (Opinion Survey) in Wright et al. (2012, 2015). Deception abilities may vary in different domains and the “deceptive-general ability” may only exist mainly on opinion topics. However, when we analysed the deceptive abilities in the Opinion Survey task separately, there was still no correlation (r = .051, p = .628).

Second, we had non-instructed lies, i.e. Choice cue in the paradigm whereas

Wright et al (2012, 2015) only had instructed lies. In the current study, on some trials participants had the choice to lie or to tell the truth, which may have affected the results. However, no correlation was found also when we excluded the Choice trials and analysed the deceptive abilities in the Opinion Survey task alone (r = .003, p = .973). With regard to this task design, we should note that the total numbers of truth and lies participants told was not even, which means that their *d*'s are based on different numbers of truths and lies judged. On the other hand, truth-bias has been widely found in deception studies (Anderson, Ansfield, & DePaulo, 1999; Levine, Park, & McCornack, 1999; McCornack, Parks, & McLaughlin, 1986), where judges tend to judge more messages as truths than lies. We also found a truth-bias in the current study (64.2%). According to Levine et al. (1999), accuracy of judgment is a positive linear function of the ratio of truthful messages to total messages. In other words, the more truthful messages the Sender told, and the more accurate the judgement by the Receiver due to truth-bias. Thus, in the current study, people's higher ability of generating lies might be due to telling more lies when given a choice, because the more lies they told and the lower the chance they would be detected. Nevertheless, the SDT we used in the current study allows lie-truth discriminability (d'_{Receiver}) to be measured independently of judgment bias. Thus, truth-bias should not affect the capacity to discriminate lies from truths. Furthermore, we calculated the ability to generate lies (d'_{Sender}) again without the 12 Choice trials and found a high positive correlation ($r = .832$, $p < .001$) between d'_{Sender} with Choice trials and d'_{Sender} without Choice trials. This shows that those who are better liars still show higher ability to generate lies, even when the Choice trials are excluded.

Third, in Wright et al (2012, 2015)'s paradigm, each participant had to make both truthful and deceptive statements on each of the opinion topics in different trials. In contrast, in our paradigm people only saw each item once and either lied or told the truth. It is unclear what impact item repetition might have had on the results. Repeating items introduces potential problems for two reasons: (i) from the sender's perspective, lying about an item one just told the truth about (and vice versa) may be different than lying about a new item, and (ii) from the receiver's perspective, having judged that a sender told the truth (or lied) about an item earlier in the game, would constrain the judgment on the same statement later on.

Fourth, the questions used in the Opinion Survey were probably not the same as those used in Wright et al (2012, 2015). These questions may vary in a number of different ways, for example in how controversial they are in the studied population. Lies told by a Sender may be more difficult to detect if they are about controversial topics than if they are about less controversial topics. For example, if a Sender lies by claiming to be in favour of a topic that is known to be unpopular in the studied population, then a Receiver can use this background information to predict that the Sender probably is lying. In the current study, participants had to rate how much they agreed with each opinion topic on a 1 (absolutely agree) to 6 (absolutely disagree) scale before the task. The 'Agreement Index (AI)' was calculated for each topic by giving each individual +1 if they rated 1, 2 or 3; -1 if they rated 4, 5 or 6, and by summing up the scores of all 93 participants. After converting the scores to their absolute value, higher AI means that most participants had a similar opinion about the topic and low AI (close to zero) means participants showed diverse opinions toward the topic. There was a strong positive correlation between AI and accuracy of judgment by receivers ($r = .908$, $p < .001$), indicating that if most people have a similar opinion about the topic, then it will be easier for receivers to detect deception. In contrast, if people do not have a consistent opinion about the topic, the accuracy of lie detection will drop. This finding shows that it is not only individual differences in lie-detection ability that matter, the topic people lie about also plays a critical role for receivers to judge. In other words, in addition to how credible the sender is, receivers also consider the topical context in making their judgment. However, since all participants lied or told the truth about the same set of topics and lie detection rates were only 60%, it is unlikely that the AI played an important role in affecting the correlation results between lie generation and detection abilities.

Our data suggest that the correlation between lie generation and detection abilities reported by Wright et al. (2012) is not a robust finding. At best, it is highly dependent on the details of task design and stimulus set; at worst, it is a false positive.

Based on the current findings, the ability of people to generate lies correlates with their creative achievement, especially in the performance domain. However, their ability to detect lies shows no relationship with creative cognition. Together with the lack of a correlation between generating and detecting lies, these results provide evidence against the idea of a general deception ability. This is also in line with a number of other findings. Kraut (1978) showed that accuracy in judging the veracity of one person is independent of accuracy in judging the veracity of another. One meta-analysis has shown that all individuals are barely able to detect lies and the real differences in detection ability are miniscule (Bond Jr & DePaulo, 2008). They suggested that in an individual differences sense, the accuracy of a deception judgment depends more on the liar than the judge. The individual differences for liars (i.e., the positive correlation with creative cognition) found in the current study also support this suggestion that the ability to generate lies of a liar plays a major role in the accuracy of a deception judgement. In contrast, the ability of detecting lies showing no correlation with both creative cognition and the ability of generating lies may indicate that there is no specific ability associated with lie detection.

Although these results are correlational, future research could also directly examine whether creative cognition influences the ability to generate lies by manipulating an independent variable (i.e., creative cognition) or vice versa. For instance, participants could be randomly assigned to two conditions, where one group of participants is primed to think creatively before the interactive deception game and the other one is a control group. That way one could examine whether participants who have been primed to think creatively show better ability of generating lies compared with the control group.

Summary and Conclusion

The main goal of this study was to address one of the aims of this thesis — to determine whether creative cognition contributes to deceptive communication using an individual differences logic. The results show that the ability to generate lies was predicted by the number of uncued lies told when given the choice and by a higher creative achievement score in the performance domain. In contrast, the ability to detect lies was not predicted by any of the creativity measures. Importantly, no hints of a correlation between the ability to generate and to detect lies were found challenging the idea of a deception-general ability. This research provides the first evidence for a link between creative cognition and deceptive abilities. Although these findings are correlational, they suggest that creative cognition is one of the factors contributing to the ability to generate believable lies.

CHAPTER 3

STUDY 2 – An ERP Study of Uncued lies in a Bluffing Game

Introduction

In the behavioural study reported in Study 1, participants delivered deceptive messages in an interactive paradigm where they were not only instructed to lie but they were also allowed to lie at will. This paradigm, involving a key component of deception – decision-making, successfully enabled participants to decide when to lie in an experimental setting. To further investigate this type of realistic and interactive deception setting in the laboratory, this event-related potential (ERP) study aimed to create an ecologically valid paradigm suitable for electroencephalogram (EEG) recording to examine the neural processes underlying spontaneous deception.

There has been increasing interest in the cognitive and neural mechanisms underlying deception. Conceptually, lying is usually more cognitively demanding than truth telling (DePaulo et al., 2003; Vrij et al., 2006). First, liars need to invent a credible story and to suppress the truth that is also actively kept in working memory. Second, one needs to monitor one's fabrication so that the fake story is plausible within the context of the interaction, otherwise the target might become suspicious. Third, in order to maintain consistency, extra cognitive effort is required to remember what was said (Christ et al., 2009; Spence et al., 2001; Vrij, Granhag, Mann, & Leal, 2011; Walczyk, Roper, Seemann, & Humphrey, 2003). Fourth, liars may attempt to control behaviours by avoiding exhibiting behaviours that may appear suspicious and also by showing behaviours to appear credible (Buller & Burgoon, 1996; Hocking & Leathers, 1980). Increased cognitive load for lying is in line with recent fMRI results showing more activation in brain areas associated with cognitive control (e.g., prefrontal cortex and anterior cingulate cortex) for lying than telling the truth (Bhatt et al., 2009; Christ et al., 2009; Langleben et al., 2002; Spence et al., 2001).

Event-related potentials (ERPs), which provide millisecond temporal resolution, have been an attractive means to study the neural dynamics of cognitive processes underlying deception. ERP results during the last decade have shown that lying has a different electrophysiological signature than truth telling. For example, the P300 component, known to be reduced by cognitive load, is significantly smaller during lying compared to truth telling (Hu et al., 2011; Johnson et al., 2003, 2005; Pfister, Foerster, & Kunde, 2014; Suchotzki et al., 2015; Wu et al., 2009). The reduced P300 for lying reflects the fact that lying requires the truth to be monitored, inhibited and kept active at the same time (Christ et al., 2009; Vrij et al., 2011). Note that the P300 component is actually enhanced in concealed information paradigms (more on this in Study 3) because in these paradigms the "oddball" recognition aspect of the task is much more prominent than the deception aspect. Another component, the frontal N200, has also been found to be relevant for lying. The enhanced N200 for lying compared with truth telling was interpreted as an indication that lying comes with increased response conflict (Hu et al., 2011; Suchotzki et al., 2015; Wu et al., 2009). Also, several ERP studies assessed the roles of strategic monitoring processes in deceptive responding by investigating response-locked ERPs (Johnson et al., 2004, 2005, 2008). The medial frontal negativity (MFN) has been found to be larger for deceptive responses than truthful responses (Johnson et al., 2008), consistent with evidence that the MFN is generated in or near anterior cingulate cortex (ACC) (Gehring & Willoughby, 2002; Johnson et al., 2004; Ullsperger & Von Cramon, 2001) and is linked to response monitoring and response conflict (Gehring & Willoughby, 2002; Johnson et al., 2004; Mathalon et al., 2003; Vidal et al., 2000). Another ERP component, the pre-response positivity (PRP), appearing around 250 -300 ms prior to the response, has been found to be smaller for self-generated lies compared with directed lies, because self-generated lies require more strategic monitoring prior to response execution (Johnson et al., 2008). The late positive component (LPC) at parietal regions provides a sensitive index of how processing resources are allocated between two simultaneously performed tasks (Johnson, 1988, 1993). In deception studies, the LPC is smaller in amplitude for deceptive than for truthful responses (Johnson et al., 2003, 2005; Johnson et al., 2008), because some resources are allocated to secondary task (i.e., responding deceptively) while holding the information of the primary task (i.e., the truth). Recently, Panasiti et al. (2014) found that lying was associated with decreased motor readiness prior to the response, as indexed by the Bereitschaftspotential (BP). The BP is an index of motor readiness to volitional movements (Shibasaki & Hallett, 2006), and is modulated among other factors by the moral value of the selected option, with instrumental dilemmas (i.e., intended means to save others but kill one individual) showing lower motor readiness (Sarlo et al., 2012). Panasiti et al. (2014) proposed that the BP could be a good marker of the moral conflict related to spontaneous deception with reduced motor readiness for lying compared to truth telling.

There has been much discussion about the ecological validity of the current neuroscience of deception due to the unrealistic paradigms used in most studies. Vrij and Ganis (2014) provided a characterization of deception as follows: ‘deception, a deliberate attempt to convince someone of something the liar believes is untrue, is a fact of everyday life.' From this definition, it follows that deception involves deliberation and an intention to convey misleading information. However, instructing participants to lie is a feature of nearly all studies in the literature, eliminating a key component of deception. In addition, deception is socially rooted and is associated with the expectation of social consequences either positive or negative (Sip et al., 2008). Here, we aim to develop an ERP task that takes into account some of the cognitive processes underlying deception in an interactive context with information of possible reward and risk of punishment.

Recent studies have begun to develop tasks to characterize the electrophysiological signature of spontaneous deception. For example, Hu, Pornpattananangkul, and Nusslock (2015) used an incentive-based coin-guessing task (adapted from an fMRI study by Greene and Paxton (2009)) in which participants predicted the outcome of an upcoming coin flip and they had the opportunity to lie for monetary incentive because in the critical condition their prediction was not recorded in advance but it was only reported after the coin flip outcome. With this design, however, it was not possible to access the participants' private predictions and compare truthful and deceptive trials directly. Pfister et al. (2014) devised a paradigm in which participants were asked to locate a knife (i.e., upper or middle compartment) for a virtual police officer. They had to announce that they would lie or tell the truth (i.e., intention response) at the beginning of the trial, and then made a truthful or deceptive response based on the intention response when the stimulus appeared. Although the experimenters found differences between truthful and deceptive trials, announcing whether to lie or tell the truth before seeing the target stimulus was far from real-life deception. As Sip et al. (2008) suggested, in order to deceive, a person must be able to weigh the advantage to be gained by deceiving against the risks, which in this study was not possible because the decision to lie was made before seeing the item to lie about. Also, there was no motivation for participants to lie. Another paradigm used in Panasiti et al. (2014) was based on a two-card (i.e., winning card and losing card) game, in which an opponent (OP) had to pick one of the two covered cards and the participants had to communicate the outcome of the choice to the OP because the OP could not see it. For each trial, participants were completely free to decide whether to lie or not to the OP and the winner took money from the other player. Participants were told that the OP would meet them at the end of the experiment. This paradigm contained social interaction and motivation elements to lie in order to gain more money, but there was no risk or punishment if participants lied unsuccessfully. In other words, participants would never be caught when they made deceptive responses and so the choice about whether to lie or tell the truth was only affected by moral factors. However in real life, deception is typically influenced by probable gains and losses (Sip et al., 2012) and it involves both greater potential risks and rewards. In an ERP study, Sun, Mai, Liu, Liu, and Luo (2011) took this characteristic of deception into consideration by using a simulated "bill-identification" task where participants were required to pick out the genuine bill picture from a set of mixed fake ones. Telling the truth would gain them a small but guaranteed monetary reward, whereas lying might lead to a larger potential gain but might also carry the risk of a double penalty if they were caught. Participants were encouraged to use spontaneous deception without being explicitly instructed to lie. However, they were advised to make approximately equal deceptive and truthful responses in this task, and so the "free-will deception" was constrained by these instructions. Another disadvantage of this task was that participants might respond arbitrarily to the bill and perform inattentively, which was addressed by adding some probe trials on which participants had to respond truthfully.

The first aim of the current study was to develop a novel task involving some social interaction with possible gains and losses that would elicit a broad set of cognitive processes during spontaneous/uncued deception. Overall, we expected to see that the neural correlates measured under this ecologically valid paradigm would show a different pattern from that typically found in studies using instructed lies. A novel task was implemented based on the main idea of a bluffing game. In this task, one out of three different cards (i.e., High-, Mid-, and Low-value) was presented to participants in random order, and they were asked to report either truthfully or deceptively the value of the card they had just seen. They were told that they were playing a real-time bluffing game with a judge sitting outside of the EEG room and their goal was to deceive the judge in order to win more money. Reporting the card truthfully would win them the value of the card, whereas lying might win double the value of the card they reported plus an extra reward if the judge believed it, but it would also carry the risk of losing the same amount of money they could win if they got caught.

This design requires the ability to make rapid decisions involving several cognitive processes, including: (i) recognizing the target and deciding that deception is a possible course of action, (ii) evaluating the advantage to be gained by deceiving against the risks of being detected, (iii) figuring out how to manipulate the receiver (i.e., the person who receives deceptive/truthful message), which might also involve moral conflict, (iv) making a decision (deceptive response or not), (v) executing the decision while monitoring actions to ensure that selected response conform to one’s plans and goals. These cognitive processes are the key components of social decision-making underlying deception, and it is crucial to consider these different components of deception if we want to determine the neural markers of deception (Sip et al., 2008). Critically, instead of being instructed to deceive the receiver, participants were free to choose when to deceive, after evaluating whether deception was feasible or advantageous based on the current situation. In addition, as argued by Sip et al. (2012) we choose to deceive because we believe that if our deception is successful we shall be better off than if we had told the truth; meanwhile managing one's reputation enhances the moral conflict that is linked to deception (Panasiti, Pavone, Merla, & Aglioti, 2011). These arguments might suggest that one's personality traits and cognitive abilities related to generating lies would affect the number of spontaneous lies made during the task. Thus, we also planned to examine whether generating more lies would be correlated with certain personality traits and cognitive abilities.

The second aim was to see whether there were neural patterns associated with successful detection of deception. Although deception can lead to possible gains, detecting deception is also important in social interactions, and it depends on one's ability to predict and infer the mental state of the other. To avoid a possible loss, it is important to judge accurately whether the other person is lying or not. To this aim, after the bluffing game, we showed participants a series of short videos of a person playing the same bluffing game and asked them to judge whether or not the person was being deceptive in order to gain more rewards. It has been consistently reported that detection of deception based on behavioural cues is not an easy job (Bond & DePaulo, 2006; DePaulo, Zuckerman, & Rosenthal, 1980; Ekman & O'sullivan, 1991; Vrij, 2000) and people have the strong tendency to judge the behaviour of others as being truthful (Levine et al., 1999; Vrij, 2000). There has been very little work on the neural correlates associated with detecting deception. In fact, to our knowledge, only two studies have investigated this topic with neural measures both using fMRI. Grèzes, Frith, and Passingham (2004) found that there was increased activation in the amygdala and rostral anterior cingulate cortex when participants judged the actions (i.e., lifting a box) concerning the real weight of the box as reflecting deceptive intention. In contrast, Sip et al. (2010) found no differences between judging the opponent's claim truthfully or deceitfully when participants were playing a two-person game. In the current lie detection task, participants judged whether the person was being truthful or deceitful based on their facial expression and oral report. We expected to find some ERP differences between correct and incorrect judgment of deception. We also expected to see that better lie detectors would show different neural patterns compared to those with lower rate of correct judgment.

Finally, the third aim was to examine whether detection of deception would show different patterns of neural activity compared with a simple perceptual judgment (e.g., detection of motion) of the same stimuli. Therefore, we implemented another task using exactly the same stimuli for participants to judge whether the person in the video was moving his/her head or not. Both detection of motion and deception require visual attention to process the stimuli, and so early ERP components such as the P1 and N1, which are frequently of interest in perception and attention studies (Woodman, 2010), were examined and compared between these two tasks. Since the detection of motion task is a low-level visual task, it was expected to elicit a larger P1 and N1 over occipital regions than the detection of deception task early in processing. In contrast, the detection of deception involving higher-level categorization processes was expected to show differences later in processing over frontal regions compared with the detection of motion task.

In sum, we aimed at creating a deception paradigm with a social context requiring a series of key cognitive processes for successful deception. Since this experimental paradigm was developed and used for EEG study for the first time, we conducted mass univariate analyses to examine the effect of ERP components that have been reported in previous deception studies as described earlier, and to identify any unexpected effects. We aimed to compare whether spontaneous deception induced by an interactive circumstance would exhibit different electrocortical activity that reported in the instructed deception literature. Based on previous ERP studies of deception and on the deception-related cognitive processes that were proposed for the ERP components found in previous studies, we expected the following: for stimulus-locked components, (1) an attenuated parietal P300 for lying compared to truth-telling because lying is associated with higher cognitive load, (2) a larger N200 for deceptive responses due to conflict detection; for response-locked components, (3) a larger MFN for lying due to increased response conflict, (4) a smaller PRP for lying because it requires strategic monitoring to be successful, (5) a smaller LPC for deceptive response reflecting how resources are allocated to process lying, and (6) reduced motor readiness (i.e., BP) for lying, possibly due to moral conflict.

We also carried out additional analyses to compare the results with those of the first behavioural study (i.e., Study 1). As participants completed both generation and detection of deception in the current ERP study, we conducted the same signal detection theory (SDT) analysis to examine the individual differences in the ability to generate and to detect lies and the correlation between these two abilities. In addition, scores for the creativity achievement questionnaire (CAQ) and the Abbreviate Torrance Test for Adults (ATTA) were also collected to examine the correlations with deceptive abilities.

Consistent with what was found in Study 1, we expected to find a positive correlation between creativity scores and the ability to generate lies but not with the ability to detect lies.

Materials and methods

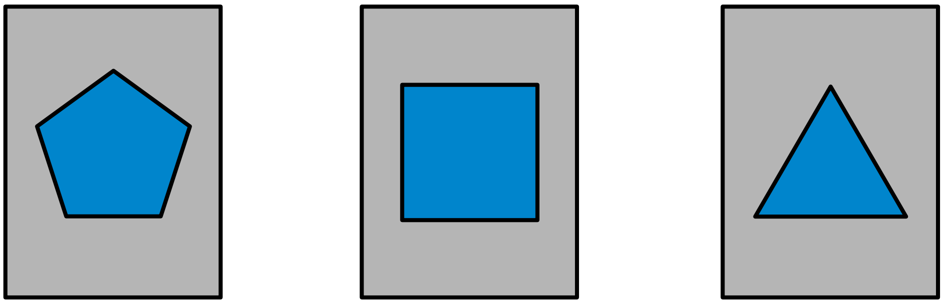
Participants

A total of 45 healthy participants (34 females and 11 males, mean age = 20.0 years, SD = 2.0), recruited from University of Plymouth, completed the head movement, bluffing game and lie detection tasks. All participants had normal or corrected vision, and no history of neurological or psychiatric disease. Data from 9 participants were excluded from the analyses (1 due to excessive muscle tension artifacts, 2 due to misunderstanding of the rules, 1 due to inattentive performance, 5 due to too few lies (< 10) to be analysed after artifact rejection), leaving 36 participants (26 females and 10 males, mean age = 20.1 years, SD = 2.2). All procedures were approved by the Plymouth University Faculty of Science and Technology Human Ethics Committee. All participants gave signed informed consent and were either given course credits or remunerated for their time in the study. The full subject information and consent form is presented in Appendix 3A.

Stimuli

In the bluffing game, the stimuli were 3 gray cards depicting the shape of a pentagon, a square or a triangle in blue against a black background (Figure 3-1). Each card had a different monetary value. Every stimulus was presented after an average inter-trial interval (ITI) of 1000 ms, ranging from 750 ms to 1250 ms. There were 180 stimuli composed of 60 High-value, 60 Mid-value and 60 Low-value cards. To counterbalance the assignment of values to shape, there were two versions of the stimuli: (i) one with more complex shapes representing higher values (pentagon: High-value; square: Mid-value; triangle: Low-value), and (ii) one with less complex shapes representing higher values (triangle: High-value; square: Mid-value; pentagon: Low-value). Each stimulus was presented until a response was made.

Figure 3-1. Card stimuli used in the bluffing game



In the head movement and lie detection tasks, the stimuli were clips produced during a pilot study of the bluffing game. 135 trials were selected as stimuli. Each trial was cut into two parts: between card onset and before the onset of the oral response, and after such onset. By doing this, we were able to know when participants heard the voice cue in each trial so that we could examine whether the voice cue played a role in detecting lies or not. There were 70 truths (High: 25; Mid: 25, Low: 20), 53 small-lies (Mid: 27; Low: 26) and 12 big-lies (small-lies and big-lies will be explained in the Procedure session). The stimulus was presented after a blank of 1000 ms and a fixation of 1000 ms. The duration of each stimulus varied, depending on the length of the clip.

Procedure

The study was divided into two sessions. The first session was the EEG experiment, where participants completed the head movement task, the bluffing game and the lie detection task. After the EEG experiment, participants were scheduled for a second session to complete cognitive tests and questionnaires.

In the first session, after setting up the EEG cap and electrodes, participants were seated on a comfortable chair, 115 cm from a computer screen in a dark room.

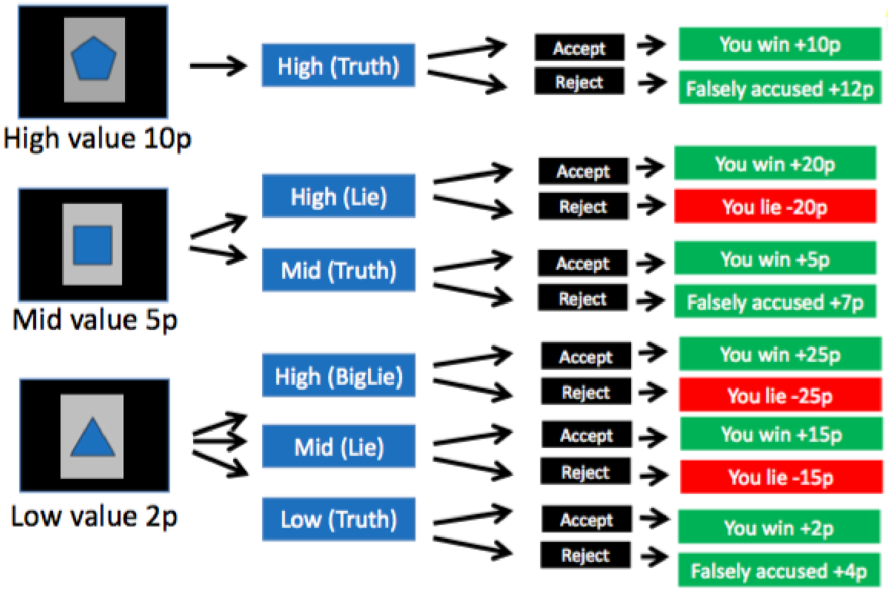
Stimuli for the three tasks were presented using Presentation® software (Version 18.0, Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com).

In the bluffing game, participants were asked to report either truthfully or deceitfully the value of the card on each trial using a 3-button response box with their dominant hand. If participants chose to report the value of the card deceitfully, they were only allowed to report a card type with a higher value (that is, they were not allowed to lie counter productively by reporting a lower value); otherwise the trial was considered as rule-breaking and a penalty occurred. Under this rule, when seeing a High-value card participants could only respond truthfully, a trial type we will refer to as “Cued truth”. When seeing a Mid-value card, participants could either report truthfully by pressing the Mid-value button (Uncued truth) or report deceitfully by pressing the High-value (Small lie) button. There were three options when participants saw a Low-value card, report truthfully by pressing the Low-value button (Uncued truth) or report deceitfully by pressing either the High-value (Big lie) or the Mid-value (Small lie) button. After pressing the button, an image of a microphone appeared and participants were asked to say aloud the value they had just reported on the response box. The response participants gave via the button had to match their verbal response. After the verbal report, participants pressed a button with their thumb on their non-dominant hand to continue.

Participants were filmed for the duration of the task and the video was streamed in real time to a laptop outside the room. Participants were told that a judge sitting outside the room would assess whether their response was truthful or not by watching the live stream. We also told participants to make sure the verbal report was clear enough for the judge to understand. However, the feedback (i.e., accept or reject) from the judge was fixed by the program to make sure everyone received the same amount of accepts and rejects. Participants won or lost money depending on the response they made and the judge's assessment of their response. If the deceptive response was accepted, participants gained the value of the card and a bonus, with bigger the lies corresponding to bigger bonuses. For example, reporting a High-value card when participants actually had a Low-value card gained a bigger bonus than reporting it as a Mid-value card. However, if the deceptive response was rejected, participants were penalised by losing the value of the card and an additional penalty that is, bigger lies corresponded to bigger penalties. On trials when a truthful response was rejected, participants gained the value of the card plus a small amount of compensation.

The gains and losses accrued throughout the task and the updated score was shown at the end of every trial. The objective was for the participants to deceive the judge in order to win as much money as possible. We told participants to appear as credible as possible, regardless of whether they told a lie or the truth. Participants were reminded that the money they won from the game would be paid in cash. The list of payoff possibilities is shown in Figure 3-2.

Figure 3-2. List of payoff possibilities in the bluffing game



The mapping of buttons to card values was counterbalanced. There were two versions: (i) button 1 referred to High-value, button 2 referred to Mid-value and button 3 referred to Low-value, (ii) button 3 referred to High-value, button 2 referred to Mid-value, and button 1 referred to Low-value. Together with the two versions of the assignment of value to shape (mentioned in the *Stimuli* section), there were 4 (2 x 2) instruction-counterbalancing versions in total.

The 180 trials were divided into 3 blocks with equal numbers of High-value, Mid-value and Low-value cards (20 of each) pseudo-randomly presented in each block with the constraint that the same card was never repeated more than 3 times in a row. The accept/reject feedback ratio was fixed at 60/40 so 36 of the trials in each block were accepted and 24 were rejected. This distribution of feedback was used in previous ERP studies (Kireev, Pakhomov, & Medvedev, 2007; Kireev, Starchenko, Pakhomov, & Medvedev, 2007) and it was shown to elicit cheating in participants. Again, the same feedback was never repeated more than 3 times in a row. Each block had a fixed (stimuli) presentation order. The order of these 3 blocks was counterbalanced (123, 231, 312), and each order also had a reverse version, (i.e., where the last stimulus was presented as the first one). The feedback also followed this fixed structure. Thus, there were 6 different fixed orders in total for both the cards and the feedback. Each participant was assigned to a version of instructions and order randomly.

Participants received a practice session before the experimental trials. There were 15 trials with the same number of each card type. The accept/reject feedback ratio was also fixed at 60/40 during this practice session. The door of the testing room was open so participants could see that the judge was making judgments by pressing a keypad connected to the laptop. By doing so, participants were induced to believe that the feedback was provided by the judge.

In the lie detection task, participants watched a series of short clips where the person in the video played the same bluffing game they had just played. Each clip was followed by a question mark and participants were asked to judge whether the person in the clip was lying or telling the truth by pressing one of two buttons on a response box. The mapping between response keys and conditions was counterbalanced across subjects. The 135 trials were divided into 2 blocks with 68 trials in the first block. There were two fixed orders of clip presentation.

In the head movement task, participants watched the same stimuli as the lie detection task, but the goal was to judge whether the person in the clip was moving his/her head or not by pressing a button on a response box. The mapping between response keys and conditions was counterbalanced across subjects. The 135 trials were divided into 2 blocks with 68 trials in the first block. There were two fixed orders of clip presentation. Participants received the same 4 practice trials before the real task for head movement and lie detection tasks.

Participants completed the 3 tasks in the same order: head movement, bluffing game and lie detection. This order was constrained by the following considerations. If participants had done the bluffing game first, before the head movement task, they would notice that the person in the clip was actually playing the bluffing game when watching the clips in the head movement task. This could produce some automatic deception detection activities in the head movement task. Moreover, if participants had done the lie detection task before the bluffing game itself, they would have trouble understanding the rules and the meaning of the verbal reports. By arranging the lie detection task after the bluffing game, participants understood how to judge the clips right away because they just carried out that task.

After completing the three tasks, participants were asked if they had any trouble understanding the rules of the bluffing game by filling out a follow-up questionnaire. The full follow-up questions are presented in Appendix 3B. All participants reported that they had no difficulty in understanding the rules and found the bluffing game interesting. However, occasionally participants did not believe that they would be awarded their winnings in cash until the reward was given.

At the end of study, participants were informed that during the bluffing game, the accept/reject feedback was driven by computer.

Cognitive tests and Questionnaires

Participants came back for the second session individually to complete two creativity tests, the Creativity Achievement Questionnaire (CAQ) (Carson et al., 2005) and the Abbreviate Torrance Test for Adults (ATTA) (Goff, 2002), and the psychometric tests, the Marlowe Crowne Social Desirability Scale (MCSD) (Crowne & Marlowe, 1960) and the Lie Acceptability Scale (Oliveira & Levine, 2008).

Creativity Achievement Questionnaire (CAQ): The CAQ is based on the sum of creative products generated by an individual throughout her/his lifetime. This CAQ self-report checklist consists 10 different areas of talents with total 96 items. Each ten standard domains or art (visual arts, music, dance, creative writing, architectural design, humor, and theatre and film) and science (culinary arts, inventions and scientific inquiry) contains eight ranked questions weighted with a score from 0 to 7. Participants are asked to place a checkmark on the description fitting their achievement best.

Abbreviated Torrance Test for Adults (ATTA): The ATTA is a shortened version of the Torrance Tests of Creative Thinking (TTCT) with three activities given 3 minutes to respond each. The ATTA provides substantial insight into the creativity of adults by quantifying verbal and figural creative strengths. The creativity index is measured by four norm-referenced abilities (i.e., fluency, originality, elaboration and flexibility) and fifteen criterion-referenced creativity indicators.

Marlowe-Crowne Social Desirability Scale (MCSD): This 33-item scale measures the tendency to produce social desirable responses. This scale was used to detect potential biases toward responding in socially desirable ways on other questionnaires administered in this study. Participants were asked to answer Truth (T) or False (F) to each item

Lie acceptability scale: This 11-item scale measures an individual's attitude about deceptive communication. People likely hold different opinions about deceptive communication. For example, people who see lying as acceptable will view deception as one viable tactic for accomplishing personal goals. Participants were asked to indicate the extent of agreement to each item from a 1 to 7 scale.

Electrophysiological data acquisition

The electroencephalogram (EEG) was sampled at 8192 Hz using a Biosemi ActiveTwo system. EEG data were collected from 32 electrodes Ag/AgCl electrodes arranged according to the 10-20 system, and loose lead electrodes below the left eye to monitor eye blinks and vertical eye movements, and on the left and right mastoids. Horizontal eye movements were monitored using 2 loose electrodes places on the outer canthi of the right and left eyes. The data were downsampled off-line to 512 Hz before further processing. For analyses, data were re-referenced off-line to the average of the two mastoids for consistency with most of the existing literature on the topic.

Behaviour analysis

Performance in the bluffing game (i.e., how good participants were in telling lies) and the lie detection task (i.e., how good they were at judging whether a person was lying or not) were analysed using Signal Detection Theory, SDT (Green & Swets, 1966). For the bluffing game, the recorded videos of each participant were analysed off-line. The discriminability of a participant's truths and lies was indexed by d'_{Sender} and the perceived credibility of a participant regardless of the veracity was indexed by $C'_{\text{Sender.}}$ For the lie detection task, d'_{Receiver} indicated the capacity of discriminating lies from truth and $C'_{\text{Receiver.}}$ indicated the tendency for a participant to endorse a given statement as truthful (credulity). With these indices, more successful lying was indicated by more negative values of d'_{Sender} and better lie detection was indicated by more positive values of d'_{Receiver}, as explained in Study 1.

ERP analyses

For the bluffing game, stimulus-locked ERPs were calculated for a 1200-ms epoch (including a 200-ms baseline), beginning 200 ms prior to stimulus onset.

Response-locked ERPs were calculated for a 1450-ms epoch (including a 200-ms baseline), extending from 1100 ms before the response until 350 ms after the response. For both the lie detection and head movement tasks, stimulus-locked ERPs were calculated for a 1200-ms epoch (including a 200-ms baseline), beginning 200 ms prior to stimulus onset. Independent component analysis (ICA) was conducted for 65% of participants using the *runica* algorithm to identify ocular and muscle artifacts (Mognon, Jovicich, Bruzzone, & Buiatti, 2011; Winkler, Haufe, & Tangermann, 2011). The results of the artefact detection were visually inspected before the components were manually removed. Subsequently, trials still contaminated by blinks, eye movements, muscle activity or amplifier blocking were rejected off-line. The data were low-pass filtered at 30 Hz.

Bluffing game

For the bluffing game, ERPs of interest were N200, P300, MFN, PRP, LPC and BP. To examine the effects of these ERP components, mass univariate analyses with powerful corrections for multiple comparisons were conducted on the difference amplitudes of the comparison between Cued truths and Small lies, the comparison between Uncued truths and Small lies, and the comparison between Cued truths and

Uncued truths. The individual *erp* files were collected to construct a *GND* variable using the Mass Univariate Toolbox function *erplab2GND*.m with default baseline settings. The differences of waveforms between conditions were calculated using the function *bin_dif*.m. In order to reduce the number of comparisons, the data were decimated and resampled to 64 Hz.

Given that the current paradigm was used for the first time, we tested all electrodes and time steps between 100 and 900 ms for stimulus-locked ERPs and -800 and 200 ms for response-locked ERPs. To correct for multiple comparisons, we used the Benjamini and Hochberg (Benjamini and Hochberg, 1995) procedure for controlling for false discovery rate (FDR) as this method is suggested to be best suited for exploratory studies (Groppe, Urbach, & Kutas, 2011a, 2011b).

To identify ERPs of interest existed from the mass univariate results, each ERP component was defined by electrodes and time windows below based on the results of previous research (Hu et al., 2015; Hu et al., 2011; Suchotzki et al., 2015) as well as on aggregated grand average from trials collapsed across conditions (Brooks, Zoumpoulaki, & Bowman, 2017). For the N200 and P300, ERP components were stimulus-locked to the onset of the card during the time window between 250 and 350 ms for the N200 and between 450 and 650 ms for P300. For the MFN, PRP, LPC and BP, ERP components were response-locked to the onset of the response at mid-frontal electrodes (Fz, F3, F4, FC1, FC2 and Cz) during the time window between 0 and 100 ms for the MFN, and -200 and 0 ms for the PRP. For the LPC, the response-locked ERPs were defined during the time window between -250 and 0 ms at mid-posterior electrodes (Pz, P3, P4, CP1 and CP2). The BP was defined between 800 ms and 400 ms prior to the response at medial frontal electrodes (Fz and Cz).

All mass univariate analyses were carried out using the Mass Univariate ERP Toolbox (http://www.openwetware.org/wiki/Mass_Univariate_ERP_Toolbox).

Repeated-measures of ANOVAs were also conducted on the mean amplitude in a priori windows to examine each ERP component. The full details of methods and results are presented as supplementary result in Appendix 3C.

Since 20 participants (out of 36) had less than 10 valid trials in the Big lie condition, the current ERP results only included Small lie condition in analyses.

Lie detection task

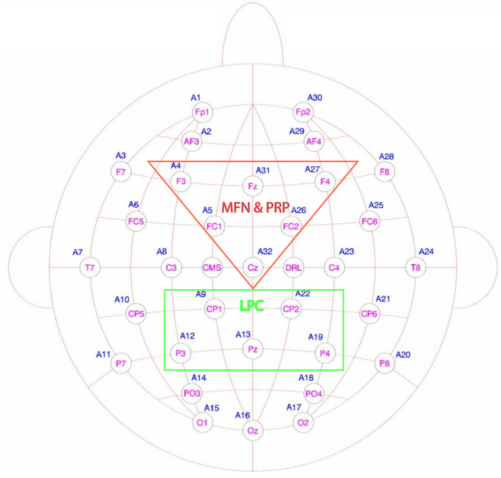
Participants were split into "Good-judge" and "Bad-judge" groups based on the median according to their behavioural performance on lie detection judgement in order to examine whether those who were better at detecting lies would show distinct electrophysiological signatures reflecting their better lie detection ability.

For the lie detection task, ERPs of interest were the P1 and the N1 (i.e., visual components). Based on aggregated grand average from trials of collapsed conditions (Brooks et al., 2017) and previous ERP literature (Hillyard & Anllo-Vento, 1998; Vogel & Luck, 2000), visual components were defined at Oz, O1, O2, Pz, PO3, PO4, P7 and P8 in the time window between 80 and 150 ms for the P1 and between 150 and 250 ms for the N1.

Mass univariate analyses were conducted on the difference amplitude of the comparison between stimulus types (i.e., truthful and untruthful event) and judgement types (i.e., judged as truth and untruthful event) across groups. The same mass univariate analyses were also conducted between Good-judge and Bad-judge groups. As we did not have prior hypotheses (apart from visual components) about which electrodes and time windows would demonstrate the effect, we tested all electrodes and time window between 100 and 900 ms.

In addition to the onset of the clip, we also locked to the onset of the speaking cue (i.e., audio onset). We performed the same mass univariate analyses on the audio onset to demonstrate the overall pattern of ERP waveforms between Good-judge and Bad-judge groups, and across groups.

Figure 3-3. Diagrams showing the sites analysed for the MFN, PRP and LPC



Head movement task

The only reason to have this task in the current study was to compare it with the Lie detection task. Therefore, the ERP data of the Head movement task are not analysed separately.

Lie detection and Head movement

To assess potential ERP differences between judging if a person was lying or telling the truth and judging a physical feature (i.e., head motion), the same mass univariate analyses were carried out as in the lie detection task but using the factor Task (detection, head movement), instead of Group.

3. Results

3.1 Bluffing game

3.1.1 *Behaviour*

Note that participants were only allowed to lie when the card type was Mid-value and Low-value. Thus, participants had 120 chances (60 Mid-value cards and 60 Low-value cards) to lie spontaneously in this task. Participants generated an average of 52.9 lies (SD = 17.2, range from 15 to 97) with 41.5 Small lies (SD = 13.6, range from 15 to 71) and 11.3 Big lies (SD = 9.6, range from 0 to 37). Thus, there was an average 44.1% chance for participants to generate a lie in this experimental setting. Participants only made an average of 0.8 mistakes (SD = 1.0, range from 0 to 4), due to breaking the rules.

We used paired t-tests to compare the Response time (RT) between the Total truth and Total lie conditions. The Total truth trials were subdivided into "Cued truth" and "Uncued truth" categories, and the Total lie trials were subdivided into "Small lie" and "Big lie" categories. For the Big lie category, 5 participants in total were excluded (4 because they did not generate any big lies and one because of a Z-score greater than 3). The descriptive statistics of these six variables are shown in Table 3-1. For the pairwise comparisons, only 4 subgroups of truth and lie (i.e., Cued truth, Uncued truth, Small lie and Total lie) were compared using paired t-tests in order to compare them with ERP results (Table 3-2). The reasons for picking these 4 subgroups for the paired t-tests are as follows. First, we assumed that Cued and Uncued truths rely on different cognitive processes, so these two subgroups were used in the comparisons instead of Total truth. Second, due to insufficient ERP data for the Big lie condition, these comparisons only involve Small lie in order to match the ERP results. Finally, Total lie were also involved in these comparisons in order to match the supplementary results. RTs for Cued truths (M = 794.6 ms, SD = 221.3 ms) were faster than for Uncued truths (M = 870.2 ms, SD = 202.6 ms), $t(35) = -3.95$, $p < .001$, $d = 0.66$. RTs for Small lies (M = 902.2 ms, SD = 195.9 ms) were slower than for both Cued truths, $t(35) = 4.87$, $p < .001$, $d = 0.81$ and Uncued truths, $t(35) = 2.58$, $p < .05$, $d = 0.43$. However, after Bonferroni correction, the difference between Small lies and Uncued truths was not significant. RTs for Total lies (M = 887.6 ms, SD = 195.4 ms) were slower than for Cued truths, $t(35) = 4.07$, $p < .001$, $d = 0.68$, but not Uncued truths, $t(35) = 1.27$, $p = .213$.

At the end of the game, participants received a payment of £12.3 (SD = 1.7) on average.

In order to compare the results with those of Study 1, performance was also analysed using Signal Detection Theory, SDT (Green & Swets, 1966). The recorded videos of each participant were analysed off-line by an experimenter. The discriminability of the person's truths and lies and the perceived credibility of a person were indicated by these two measures respectively: $M\ d'_{\text{Sender}} = 0.323$, SD = 0.293; $M\ C_{\text{Sender}} = -0.165$, SD = 0.126 (these results were based one 26 participants as 10 participants' data were lost due to technical reasons). More successful deception is indicated by more negative values of d'_{Sender} and higher credibility regardless of the veracity is indicated by more negative values of C_{Sender}.

Table 3-1. Mean (SD) response time

	Mean	SD
Total Truth	835.2	206.3
Cued Truth	794.6	221.3
Uncued Truth	870.2	202.6
Total Lie	887.6	195.4
Small Lie	902.2	195.9
Big Lie	799.6	219.6

Note: "Total Truth" is the combination of "Cued Truth" and "Uncued Truth"; "Total Lie" is the combination of "Small Lie" and "Big Lie".

Table 3-2. Paired t-tests of response time (RT)

	Cued Truth	Uncued Truth	Small Lie	Total Lie
Cued Truth				
Uncued Truth	**$t(35) = -3.95$ ($p < .001$)**			
Small Lie	**$t(35) = -4.87$ ($p < .001$)**	[a]$t(35) = -2.58$ ($p < .05$)		
Total Lie	**$t(35) = -4.07$ ($p < .001$)**	$t(35) = -1.27$ ($p = .213$)	**$t(35) = 3.61$ ($p < .001$)**	

Six comparisons were conducted using Bonferroni adjusted alpha level of .008 (.05/6) per test. [a] Using Bonferroni adjusted alpha level, the comparison between Uncued truth and Small lie conditions is not significant

3.1.2 *Stimulus-locked ERPs*

3.1.2a Cued truth vs. Small lie

A mass univariate analysis with FDR control was conducted on the difference between Cued truths and Small lies. A raster diagram illustrates significant differences between ERPs at every electrode and time point from 100 to 900 ms (Figure 3-4a). The most pronounced ERP difference was a P300 effect from 470 to 660 ms at central and posterior electrodes. Consistent with this, the grand-average waveforms revealed a P300 peaking around 550 ms (Figure 3-4c). The grand-average waveforms revealed a N200 peaking around 300 ms at fronto-central electrodes (Figure 3-4b), but no N200 difference was found between Cued truth and Small lie conditions (Figure 3-4a). Two unexpected effects were apparent in the raster diagram (Figure 3-4a): a left frontal effect around 130 ms, and an effect at P7 around 200 ms.

3.1.2b Uncued truth vs. Small lie

A mass univariate analysis with FDR control was conducted on the difference between Uncued truth and Small lie conditions, but no significant difference between ERPs at any electrode and time point were found.

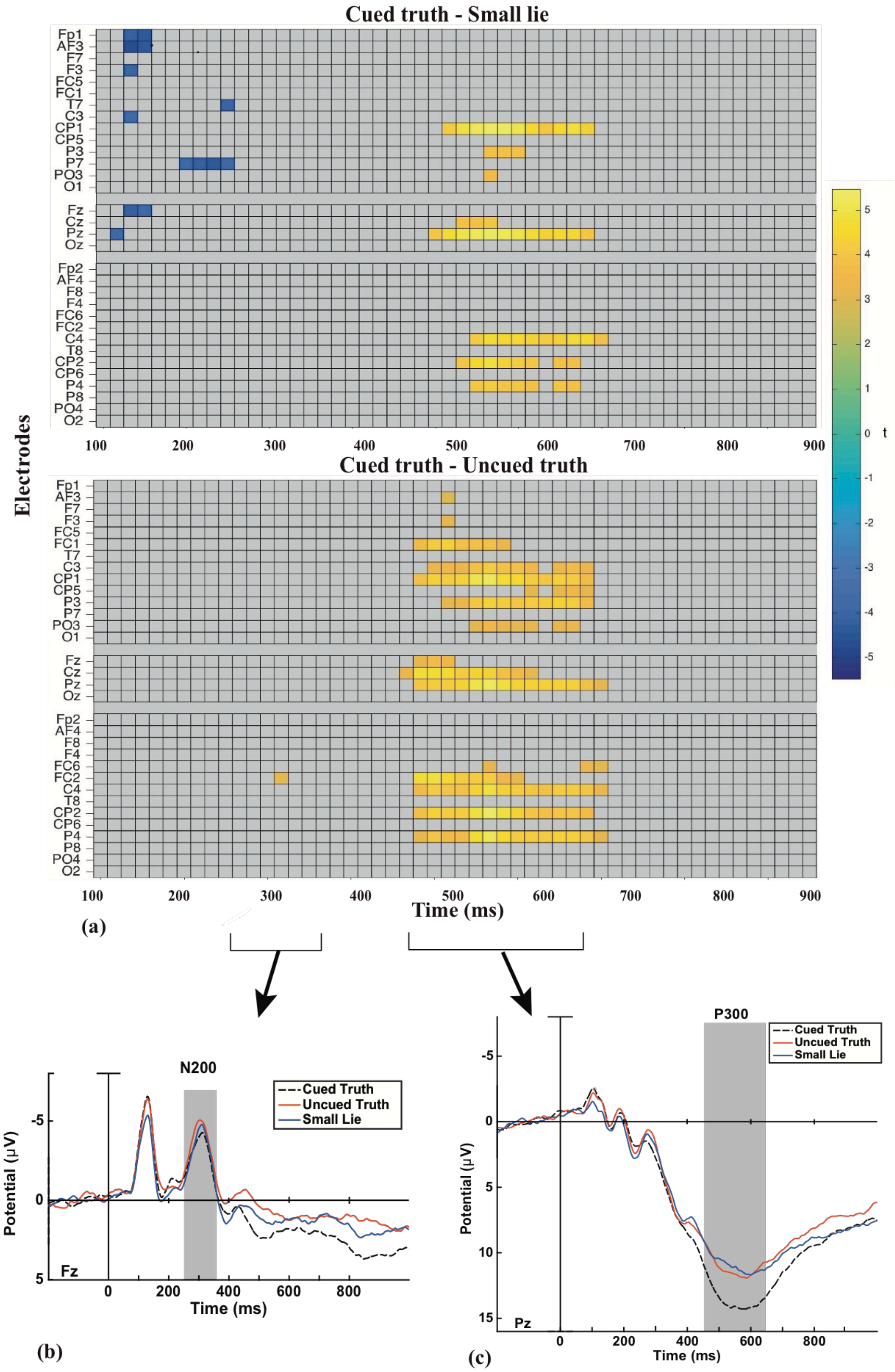
3.1.2c Cued truth vs. Uncued truth

A mass univariate analysis with FDR control was conducted on the difference between Cued truth and Uncued truth. A raster diagram illustrates significant differences between ERPs at every electrode and time point from 100 to 900 ms (Figure 3-4a). The most pronounced ERP difference was a P300 effect between 440 and 660 ms at central and posterior electrodes. No N200 effect was found.

In summary, the P300 effect was most pronounced in the comparisons between

Cued truth and Small lie/Uncued truth conditions. Neither a P300 nor a N200 effect was found between Uncued truth and Small lie conditions.

Figure 3-4. (a) Raster diagrams (Stimulus-locked) illustrating significant differences between ERPs to Cued truth and Small lie and between ERPs to Cued truth and Uncued truth. Note that there was no significant difference between ERPs to Unuced truth and Small lie so it is not shown here. Grand-average ERPs for the Cued truth, Uncued truth and Small lie condition at Fz with the N200 time window shown in grey (b) and at Pz with the P300 time window shown in grey (c)


Cued truth - Small lie
Cued truth - Uncued truth
Electrodes
Fp1
AF3
F7
F3
FC5
FC1
T7
C3
CP1
CP5
P3
P7
PO3
O1
Fz
Cz
Pz
Oz
Fp2
AF4
F8
F4
FC6
FC2
C4
T8
CP2
CP6
P4
P8
PO4
O2
100
200
300
400
500
600
700
800
900
Time (ms)
5
4
3
2
1
0
-1
-2
-3
-4
-5
t
(a)
N200
Cued Truth
Uncued Truth
Small Lie
Potential (µV)
-5
0
5
Fz
0
200
400
600
800
Time (ms)
(b)
P300
Cued Truth
Uncued Truth
Small Lie
Potential (µV)
-5
0
5
10
15
Pz
0
200
400
600
800
Time (ms)
(c)

3.1.2 *Response-locked ERPs*

3.1.2a Cued truth vs. Small lie

A mass univariate analysis with FDR control was conducted on the difference amplitude of the comparison between Cued truths and Small lies. A raster diagram illustrates significant differences between ERPs at every electrode and time point from -800 to 200 ms (Figure 3-5a). There was a pronounced LPC effect from -190 to 90 ms at all mid-posterior electrodes we focused on (i.e., Pz, P3, P4, CP1 and CP2) with a reduced LPC for lying compared to cued truths. The grand-average waveforms also revealed a LPC peaking around -100 ms (Figure 3-5d). There was a PRP effect starting from -170 ms before the response at fronto-central electrodes (i.e., Cz, FC1 and FC2) with a smaller PRP for lying compared to cued truths. An MFN effect was also found at Cz and FC2 after response. The grand-average waveforms also revealed the PRP before the response and MFN after the response (Figure 3-5c). Finally, smaller BP amplitude for lying was found compared to cued truths from -600 to -300 ms at Cz and -490 to -330 ms at Fz. The grand-average waveforms are presented in Figure 3-5b.

Unexpected effects were apparent in the raster diagram (Figure 3-5a): a slow wave effect from -730 ms to -300 ms mostly at mid-posterior and some frontal electrodes, and a frontal effect in the early time window (-800 to -700 ms).

3.1.3b Uncued truth vs. Small lie

A mass univariate analysis with FDR control was conducted on the difference amplitude of the comparison between Uncued truths and Small lies, but no significant differences between ERPs were found at any electrode or time point between -800 to 200 ms.

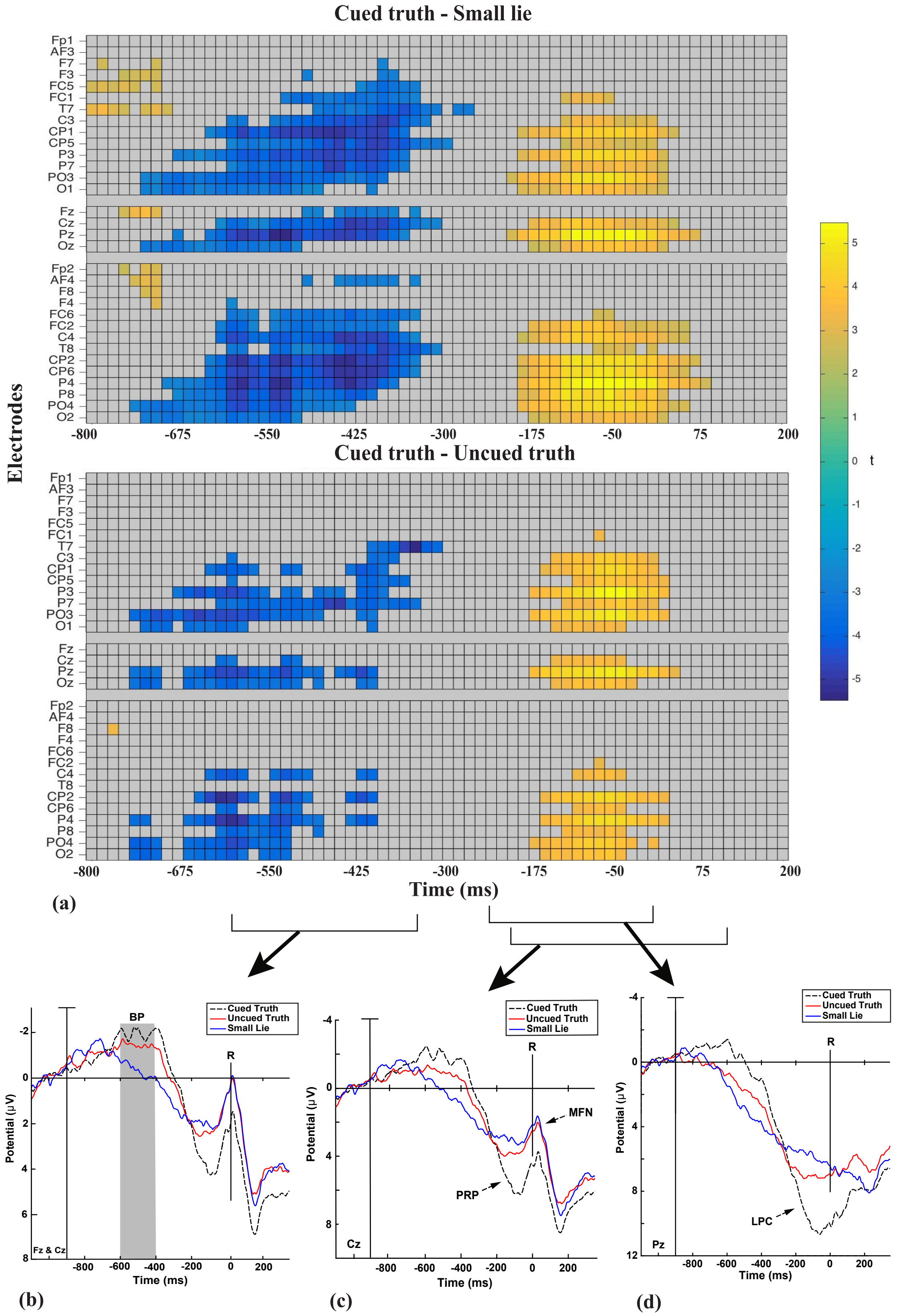
3.1.3c Cued truth vs. Uncued truth

A mass univariate analysis with FDR control was conducted on the difference amplitude of the comparison between Cued truths and Uncued truths. A raster diagram illustrates significant differences between ERPs at every electrode and time point from -800 to 200 ms (Figure 3-5a). There was a pronounced LPC effect from -160 to 40 ms at all mid-posterior electrodes we focused on (i.e., Pz, P3, P4, CP1 and CP2) with a reduced LPC for uncued truths compared to cued truths. There was a PRP effect starting from -130 ms before the response at fronto-central electrodes (i.e., Cz, FC1 and FC2) with a smaller PRP for uncued truths compared to cued truths. No MFN effect was found. Finally, smaller but sporadic BP amplitude for uncued truths was found compared to cued truths at around -600 and -500 ms at Cz only.

An unexpected effect was apparent in the raster diagram (Figure 3-5a): a slow wave effect from -740 ms to -330 ms at most mid-posterior electrodes. Compared with the slow wave effect between Cued truth and Small lies, the slow wave effect between Cued truths and Uncued truths was less pronounced.

In summary, the LPC effect was most pronounced in the comparisons between Cued truth and Small lie/Uncued truth conditions. A moderate effect of PRP was also found in the comparisons between Cued truth and Small lie/Uncued truth conditions. Again, none of the effect was found between Uncued truth and Small lie conditions.

Figure 3-5. (a) Raster diagrams (Response-locked) illustrating significant differences between ERPs to Cued truth and Small lie and between ERPs to Cued truth and Uncued truth. Note that there was no significant difference between ERPs to Unuced truth and Small lie so it is not shown here. Grand-average ERPs for the Cued truth, Uncued truth and Small lie condition at Cz (c) and at Pz (d). Another Grand-average ERPs at Fz and Cz with BP time window shown in grey (b)


Cued truth - Small lie
Cued truth - Uncued truth
Electrodes
Fp1
AF3
F7
F3
FC5
FC1
T7
C3
CP1
CP5
P3
P7
PO3
O1
Fz
Cz
Pz
Oz
Fp2
AF4
F8
F4
FC6
FC2
C4
T8
CP2
CP6
P4
P8
PO4
O2
-800
-675
-550
-425
-300
-175
-50
75
200
Time (ms)
t
5
4
3
2
1
0
-1
-2
-3
-4
-5
(a)
Cued Truth
Uncued Truth
Small Lie
BP
R
Potential (μV)
Fz & Cz
Time (ms)
(b)
Cz
MFN
PRP
(c)
Pz
LPC
(d)

3.2 Lie Detection task

3.2.1 *Behaviour*

Overall accuracy was 53.6% (SD = 5.7%), which is significantly greater than chance ($t(35) = 3.779$, $p < .001$, $d = 0.63$). Fractional rates addressing accuracy for different types of response showed a significantly higher mean accuracy for truths (M = 61.3%, SD = 9.4%) than for lies (M = 45.4%, SD = 12.6%, $t(35) = 4.95$, $p < .001$, $d = 0.83$). The percentage of statements of all types classified as truthful (truth bias) was 58.0% (SD = 9.4%), significantly greater than chance ($t(35) = 5.060$, $p < .001$, $d = 0.84$).

Performance was also analysed using Signal Detection Theory, SDT (Green & Swets, 1966) in order to compare the results with those of Study 1. The ability of discriminating lies from truths and the tendency to judge statements as truthful was quantified by $M\ d'_{\text{Receiver}} = 0.184$, SD = 0.297 and $M\ C_{\text{Receiver}} = 0.208$, SD = 0.251, respectively.

A paired t-test showed no significant differences between RTs on trials that were judged as truthful (M = 652.1 ms, SD = 301.8 ms) and deceptive (624.0 ms, SD = 259.9 ms), $t(35) = 3.779$, $p = 0.168$.

3.2.2 *ERP data*

Clip onset

Mass univariate analyses were conducted on the difference amplitude of the comparison between Good-judge and Bad-judge in 2 types of stimuli (i.e., truthful and untruthful events) and 4 types of judgements (i.e., truthful event judged as truth, truthful event judged as untruth, untruthful event judged as truth, and untruthful event judged as untruth). No significant effects were found.

Audio onset

The same mass univariate analyses were conducted on the difference amplitude of the comparison between Good-judge and Bad-judge in 2 types of stimuli and 4 types of judgements. Again, there were no significant effects.

3.3 Head movement task

3.3.1 *Behaviour*

The accuracy rate was 84.3% (SD = 4.7%), which is significantly greater than chance ($t(35) = 43.356$, $p < .001$, $d = 7.23$). Paired t-test comparing RTs between correct ($M = 522.6$ ms, SD = 144.8 ms) and incorrect ($M = 546.3$ ms, SD = 232.7 ms) trials showed no significant differences, *t*(35) = -0.935, *p* = 0.356.

3.4 Comparison between the lie detection and head movement tasks

3.4.1 *Behaviour*

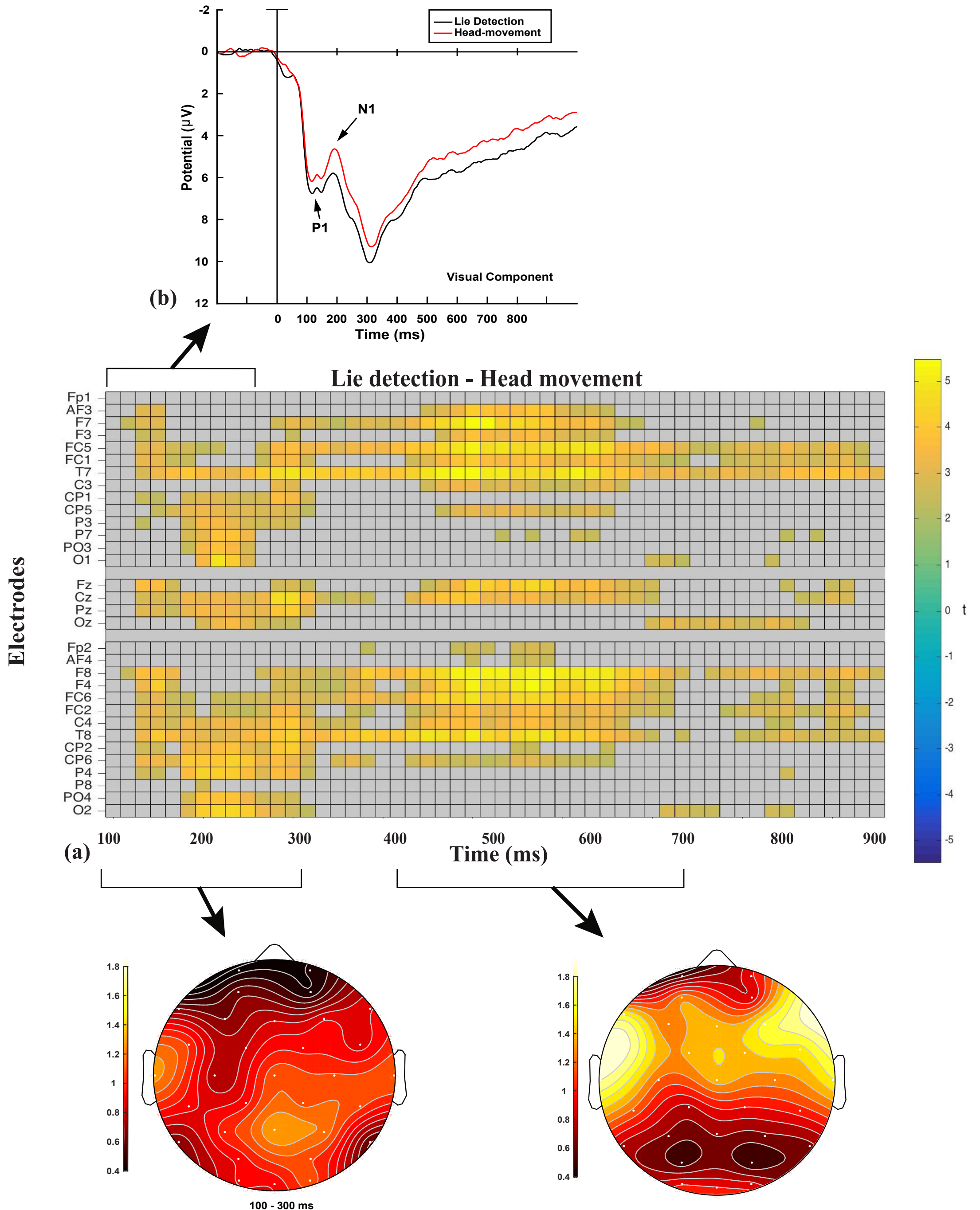
A paired t-test was used to compare the RT between the lie-detection and head movement tasks. We averaged RTs regardless of the types of judgement in each task and compared the means. RTs were slower for the lie-detection (*M* = 638.0, SD = 275.2 ms) than the head movement (*M*= 534.5 ms, SD = 178.2 ms) task, *t*(35) = 3.110, *p* < .005, *d* = 0.52. The stimuli were the same between these two tasks so this direct comparison is not affected by stimulus differences, but only by the different cognitive processes engaged by the two tasks. The longer RTs for lie-detection indicated that this task required more effort for participants to judge if a person was lying than to judge the physical feature of head movement.

3.4.2 *ERP data*

Clip onset

A mass univariate analysis was conducted on the difference amplitude of the comparison between lie detection and head movement tasks. A raster diagram illustrates significant differences between ERPs at every electrode and time point from 100 to 900 ms (Figure 3-6a). There was a N1 effect from 180 to 280 ms at occipital electrodes and grand-average waveforms are presented in Figure 3-6b. The overall amplitude for lie detection task was consistently larger than head movement task and the difference of amplitudes between these two tasks shifted from posterior to anterior regions along the epoch time.

Figure 3-6 (a) Raster diagram (Stimulus-locked) illustrating significant differences between ERPs to Lie detection and Head movement tasks. (b) Grand-average ERPs for the Lie detection and Head movement tasks at visual component (i.e., mean amplitude of Oz, O1, O2, Pz, PO3, PO4, P7 and P8). Topographic maps showing the difference between the Lie detection and Head movement tasks between (c) 100 and 300 ms and (d) between 400 and 700 ms (clip onset, collapsing stimuli)



Audio onset

The same mass univariate analysis was conducted on the difference amplitude of the comparison between lie detection and head movement tasks. None of any effect was found.

3.5 Correlations

3.5.1 *Correlations between d'_{Sender}, $d'_{Receiver}$, creativity scores and task behaviours*

Consistent with the result found in the Study 1, no correlation between d'_{Sender} and $d'_{Receiver}$ was found ($r = -.12$, $p = .559$, see Figure 3-7). Furthermore, neither d'_{Sender} nor $d'_{Receiver}$ were correlated with creativity scores (i.e., CAQ and ATTA). In addition, neither d'_{Sender} or $d'_{Receiver}$ were correlated with the number of lies generated during the task. All the statistics are shown in Table 3-3.

Table 3-3

Correlations between d'_{Sender}, $d'_{Receiver}$, creativity scores (i.e., CAQ and ATTA) and the number of Total Lie

	d'_{Sender}	$d'_{Receiver}$	CAQ-Expression	CAQ-Performance	CAQ-Science	ATTA	Total Lie
d'_{Sender}		$r = -0.12$ ($p = 0.559$)	$r = 0.203$ ($p = 0.320$)	$r = -0.256$ ($p = 0.208$)	$r = 0.181$ ($p = 0.376$)	$r = -0.075$ ($p = 0.715$)	$r = -0.248$ ($p = 0.221$)
$d'_{Receiver}$			$r = -0.276$ ($p = 0.104$)	$r = -0.073$ ($p = 0.673$)	$r = 0.136$ ($p = 0.428$)	$r = 0.162$ ($p = 0.345$)	$r = -0.069$ ($p = 0.688$)

3.5.2 *Correlations between task behaviours and cognitive questionnaires*

We performed correlational analyses between the number of lies participants generated in the bluffing game and cognitive questionnaires (i.e., MCSD and Lie Acceptability Scale). The number of small lies generated during the task was positively correlated with Lie acceptability ($r = .355$, $p = .034$) (Figure 3-8), indicating that those with higher lie-acceptability chose to generate more spontaneous lies during the task.

Figure 3-7. Scatter plot between d'_{Sender} and $d'_{Receiver}$

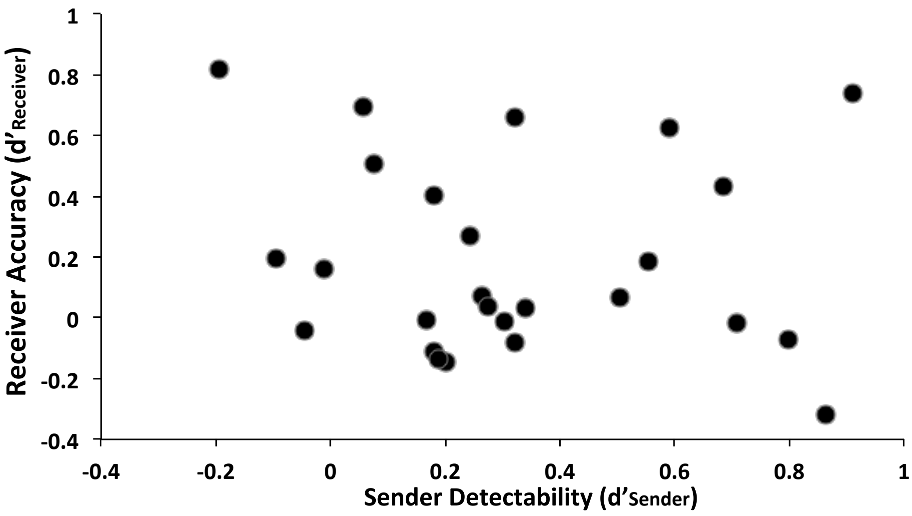
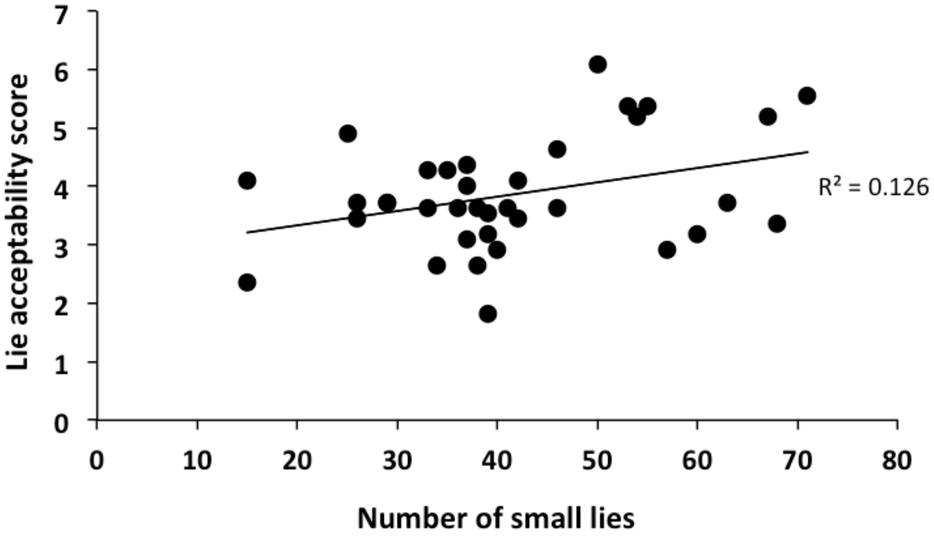


Figure 3-8. Scatter plot between the amount of Small lies and lie acceptability score



Discussion

The present study investigated the behavioural and neurophysiological signatures underlying deception by using a novel bluffing card game in which participants chose to generate spontaneous lies in order to maximize gains. The main results can be summarized as follows. First, lies showed longer RTs than cued truths but not uncued truths. Second, unexpectedly, no effect was found between lies and uncued truth. In contrast, compared with cued truths, lies elicited a smaller P300, a larger MFN and a smaller PRP, LPC and BP. A similar pattern was found comparing cued and uncued truths. Third, good and bad lie-detectors showed no ERP differences in any kind of stimulus and judgement types. Finally, comparing the head movement and lie detection tasks showed early occipito-parietal and late fronto-central effects.

Bluffing game

An advantage of the temporal resolution offered by ERPs is that it allowed us to observe multiple cognitive processes associated with spontaneous decision-making over time. The ERP results suggest that spontaneous deception is not different from spontaneous/uncued truth. Previous studies using cued conditions have reported a smaller P300 for lying compared to truth-telling in stimulus-locked analyses and interpreted it as an indication that lying is associated with higher cognitive load in order to keep the truth active, monitored and inhibited at the same time (Hu et al., 2011; Johnson et al., 2003, 2005; Pfister et al., 2014; Suchotzki et al., 2015; Wu et al., 2009). However, no such effect was found in the current study. Panasiti et al. (2014) proposed that instructed and spontaneous lying rely on different cognitive processes. Instructed lying resembles a typical executive function task where participants have to inhibit an automatic response and reverse it, whereas spontaneous lying is akin to a decision-making task where participants have to decide between two options. The lack of a P300 difference in stimulus-locked ERPs in the current study supports this notion that spontaneous deception engages decision-making processes, rather than the processes involved in reversing an automatic response.

The LPC is a positive-going ERP component generally found over parietal sites, and its amplitude has been considered a sensitive measure of processing resource allocation between two simultaneously performed tasks (Johnson, 1988, 1993). Previous studies using cued conditions have also reported a smaller LPC for lying compared to truth-telling in response-locked analyses (Johnson et al., 2003, 2005; Johnson et al., 2008) and interpreted it as an indication that lying allocates some resources to respond deceptively while holding the information of the truth. However, no LPC effect was found between spontaneous lying and truth-telling in the current study. It is likely that spontaneous truth-telling also requires allocating some resources to remember which response was made in preparation for the next trial. This strategic monitoring process to keep track of the pattern of responses during the task has been discussed in Johnson et al. (2003)'s study. In one condition of their study called "random deceptive", participants were instructed to lie randomly for half of the stimuli. There were no differences in LPC amplitude between random deceptive and random truthful responses because both were engaged in the strategic monitoring. Although participants were not instructed to lie randomly for half of the stimuli in the current study, they still needed to keep track of the pattern of responses to distribute lies and truths during the task in order to maximize gains. The lack of a LPC difference in response-locked ERPs in the current study suggests that spontaneous decision-making requires strategically monitoring one's responses regardless of whether the outcome is a lie or the truth.

In addition, Johnson et al. (2003) also directly compared random deceptive responses (i.e., participants instructed to lie randomly for half of the stimuli) and consistent deceptive responses (i.e., participants instructed to respond the opposite of the truth for all stimuli), and found much smaller P300s and LPCs in the random deceptive condition than in the consistent deceptive condition. This random deceptive condition simulating more realistic deception scenarios compared to the consistent deceptive condition provided support for the idea that real-world deception requires more cognitive resources than instructed deception measured in the laboratory.

In contrast to previous findings of a larger frontal N200 during instructed deception compared with truth-telling (Hu et al., 2011; Suchotzki et al., 2015; Wu et al., 2009), no such effect was found in the current study. The frontal N200 has been hypothesized to be involved in cognitive control for strategic monitoring, especially in conflict detection (Folstein & Van Petten, 2008; Van Veen & Carter, 2002). In the previous deception studies, the enhanced N200 for lying compared to truth-telling has been interpreted as indicating that lying comes with increased response conflict (Hu et al., 2011; Suchotzki et al., 2015; Wu et al., 2009). A possible explanation for the absence of a frontal N200 effect is proposed here. In the current paradigm, participants had to first recognize the stimulus/target so they would know whether deception was a possible course of action or not. We propose that the N200 component might reflect a target-recognition stage in which participants saw the card and distinguished the type of card based on the rules they just learned. This recognition phase had to be carried out in the beginning of the trial regardless of card type. Therefore, there was no N200 effect in any comparisons.

Regarding the MFN, we expected a larger MFN after deceptive responses than after truthful responses. However, there was no such effect when comparing lying with the uncued truth. The MFN is believed to reflect executive processes involved in conflict detection and response monitoring (Johnson et al., 2004; Ullsperger & Von Cramon, 2001). Some deception studies have found that lying elicited a larger MFN (also referred to as Correct Response Negativity, CRN) compared to truth-telling, reflecting stronger conflict for non-default answers, i.e., deceptive responses (Johnson et al., 2004, 2005; Johnson et al., 2008; Kireev et al., 2008). However, task characteristics may determine whether a truthful response is still the default (i.e., "correct" response). Verschuere and Shalvi (2014) proposed that lying may be an automatic reaction when it yields important self-profit. In other words, by rewarding and promoting successful deception, the deceptive response may be treated as the "correct" response; while the truthful response may turn into the "incorrect" response as telling the truth leads to no extra reward. Actually, Suchotzki, et al. (2015) even found a reversed effect with lying showing a smaller MFN compared with truth telling if participants were promised an extra reward if they succeeded in hiding the guilty information. In the current study, participants were motivated to lie, so it was likely that deceptive response was perceived as the "correct" response. However, the result showed no reversed effect either. A possible explanation could be that lying also carried the risk of getting caught in our paradigm but not the one in Suchotzki et al. (2015). In this case, lying was not a perfectly correct response; thus it might be the reason that neither a typical nor a reversed effect was found. These hypotheses could be empirically examined in the future studies by manipulating participants' motivation to deceive and feedback (i.e., risk of being caught).

The BP is a marker of volitional movement modulated by lower- and higher-order cognitive factors (Shibasaki & Hallett, 2006). In the current study, it is unlikely to be caused by the difference in lower-level motor activity as the same movement (i.e., press one button) was produced, and the corresponding buttons were counterbalanced. We inferred that the reduction in BP is related to high-order cognitive factors, such as decision-making based on the mental computation of the trade-off between morality and reward suggested by Panasiti et al. (2014). They found a reduction in BP for self-determined lies compared to self-determined truths and attributed this reduced motor readiness to moral conflict as lying took the reward from the opponent in their paradigm. Thus, we expected a reduced BP for lying compared to both cued truth and uncued truth in the current study. However, only the comparison between lying and cued truth showed this effect and a small BP effect was found between uncued truth and cued truth. One possible reason for this is that the reduced motor readiness might due to other cognitive factors, such as risk-taking or reward expectation. In the paradigm used by Panasiti et al. (2014), participants received the same monetary reward for both self-gain lies and self-gain truths, and they did not risk being punished. Thus, they claimed that the reduced BP was contributed by moral conflicts instead of rewards and punishments. In the current study, it was likely that the reduced BP might not only reflect deception-related moral conflict, but also involve reward expectation as uncued truths could result in a higher reward for compensation if it was a false alarm. Therefore, both lying and uncued truths showed a reduced BP compared with cued truth and no such effect was found between lying and uncued truth.

An apparent but unexpected slow wave effect was found starting from -740 to -340 ms before response at mid-posterior electrodes. This effect was shown in both comparisons between lying and cued truth, and between uncued truth and cued truth. To our knowledge, the more positive slow wave effect before response for lying compared with truth telling has not been discussed in the deception literature. This positive slow wave before the LPC observed here may be related to the amount of information required from working memory to reach a decision. In the current paradigm, both lying and uncued truth trials provided participants two choices (i.e., lying or telling the truth) where they had to retrieve information from working memory about the possible reward and punishment corresponding to the response they were going to make. In contrast, cued truth trials only allowed a choice, i.e., truth telling, so the retrieval of information was not required. Therefore, lying and uncued truths showed a more positive slow wave effect than cued truths. On the other hand, the previous literature has discussed the positive slow wave after the P300 peak in the stimulus-locked ERPs. This positive slow wave has been recognized as a response-related component. For example, it has been related to response difficulty (Kok & De Jong, 1980; Roth, Ford, & Kopell, 1978), response selection (Falkenstein, Hohnsbein, & Hoormann, 1994; Naylor, Halliday, Callaway, Yano, & Walton, 1987), and the amount of processing required for a decision (García-Larrea & Cézanne-Bert, 1998; Ruchkin, Munson, & Sutton, 1982). Nevertheless, the current results did not match directly those found in the existing literature and future work will be needed to shed further light on this slow wave.

Comparison with fMRI studies

Recently, several neuroimaging experiments using fMRI have investigated deception in more ecologically valid paradigms. In a study conducted by Sip et al. (2010), one deception game ‘Meyer,' widely played in Denmark, was developed into a laboratory version and suitable for the scanning environment. Participants played with a Confederate and they decided when to deceive the Confederate and also had to reject or accept the confederate's claim. They found that both false and truthful claims were associated with activity in the frontopolar cortex, BA10, which is considered to be at the top of a hierarchy of executive functions and associated with internally determined behaviours (Koechlin & Hyafil, 2007; Koechlin, Ody, & Kouneiher, 2003; Koechlin & Summerfield, 2007), i.e., behaviours that are not determined by external stimuli. These cognitive processes were engaged when the participants made choices in their paradigm.

This finding was different from what most previous studies had found. Deceptive-related activity had usually been located in DLPFC and VLPFC in most studies where participants were instructed to respond deceptively. These DLPFC and VLPFC regions are lower in the executive hierarchy than BA10 and are associated among other things with the selection of appropriate responses (Frith, 2000). This is consistent with traditional task designs in which participants select appropriate responses, either truthful or deceptive according to the instructed cues. These combined results provided evidence that instructed and spontaneous lying rely on disparate cognitive processes.

A recent meta-analysis of fMRI deception research directly compared neural activation patterns between social-interactive and non-interactive deception studies (Lisofsky et al., 2014). They found higher activation in the dorsal ACC, the right temporo-parietal junction (TPJ)/angular gyrus and the bilateral temporal pole (TP) in social-interactive than in non-interactive deception. The latter two brain regions have been linked to perspective taking, theory of mind and moral reasoning processes (see Mar, 2011; Saxe & Kanwisher, 2003; Saxe et al., 2006) demonstrating the importance of these socio-cognitive processes in deception. In addition, they also conducted another meta-analysis to compare the neural activation between volitional (i.e., uncued/spontaneous) and instructed (i.e., cued) deception paradigms and showed increased activation in the bilateral IPL in volitional deception. These two meta-analyses indicated that the two important components of deception, i.e., social interaction and volitional/uncued decision, are distinct underlying processes. This review study also pointed out the importance of taking these factors into account when developing new experimental paradigms.

Cued truth vs. Uncued truth

Because of the bluffing game rule, participants had no choice but to respond truthfully to the High-value card, which was different from truthful responses when they could choose at will. The direct comparison between these two kinds of truths can give us a hint about the neural activity underlying spontaneous decision-making process and also indirectly provide evidence that instructed and spontaneous deception rely on different cognitive processes. Although both cued and uncued truths are instances of truth-telling, ERPs to uncued truths showed smaller P300, PRP and LPC components than to cued truths. Modulation of these components is often reported in the comparison between instructed lies and truths. These results have a number of implications described below. First, decision-making processes themselves regardless of the type of response (i.e., deceptive or truthful) increase demands on executive control as indexed by a smaller P300 (Chen et al., 2008; García-Larrea & Cézanne-Bert, 1998; A. Kok, 2001) compared to executing specific actions (e.g., cued truth). Consistent with this idea, Hu et al. (2015) also found a smaller P300 for the opportunity condition (i.e., when the individual is free to make a truthful or untruthful response) than for the non-opportunity condition regardless of the response; the reduced P300 was proposed to reflect the engagement of executive control processes to resolve the conflict between two competing response tendencies. Second, making a choice between lying and truth-telling seems more demanding than just lying itself. From the previous literature, it is commonly believed that lying engages executive functions more (Spence et al., 2001) than telling the truth and earlier studies on instructed lying have reported longer RTs for lying (Abe et al., 2007; Abe et al., 2006; Dong et al., 2010; Johnson et al., 2003, 2005; Mameli et al., 2010; Spence et al., 2001; Wu et al., 2009). However, recent studies on spontaneous lie using ecological paradigms have shown no difference in RTs between lying and truth-telling (Karim et al., 2010; Panasiti et al., 2014; Sip et al., 2010; Sun et al., 2011) similarly to what we found in the current study, i.e., there was no RT difference between uncued truth and lie, but uncued truth showed longer RTs than cued truth. Finally, these results support the idea that there is no "specific deceptive response" based on the ERPs that can be used to investigate deception; the effects found in deception conditions are typically due to general purpose cognitive processes underlying deception instead (e.g., Ben-Shakhar & Furedy, 2012). In other words, there is no distinct neural activity pattern exclusively for deception and so detecting deception based on neuroimaging measures could provide incorrect information. The thesis will examine the vulnerability of these neuroimaging measures in Studies 4 and 5.

Novel Design

Our paradigm is based on the idea that deception is the outcome of social decision-making where people face the dilemma of deciding whether to deceive others to gain a reward or to avoid the risk of being punished by telling the truth. One novel feature of the current paradigm is that there are not just two choices (lie or tell the truth), because participants could decide the level of deception they wanted to engage in. One of the key elements of successful deception is to keep track of the false beliefs being engendered in the receiver (Sip et al., 2008). A good liar should consider that if there are too many big lies (i.e., always reporting High-value card), it might arouse suspicion in the receiver (i.e., judge) and it would increase the risk of getting caught. Keeping track of all this requires maintaining the information in working memory in order to deceive the other successfully. This feature of our paradigm makes this experimental setting closer to real life situations than other paradigms in the literature. Furthermore, this paradigm simulates a game where there is also a rule to follow to avoid incurring into a penalty for breaking it. Consequently, participants have to pay attention to the game all the time without the need to intersperse target trials during which participants have to respond truthfully in order to make sure they pay attention to the task. For example, the bill-identifying task mentioned earlier and used in Sun et al. (2011) allowed participants to lie at will. However, the lack of a rule allowed participants to just give random responses. To make sure participants paid attention, they had to insert target trials for participants to always respond truthfully. Also, because the current paradigm simulated a game, most of the participants found the game very interesting and were able to concentrate on it all the time through the experiment, at least judging from the follow-up questionnaire.

However, the closer the paradigm gets to real life in the laboratory and the more experimental control we have to sacrifice. Specifically, we could not control the proportion of deceptive trials as the number of lies varied from participant to participant. For some participants there were not enough deceptive trials for ERP analyses, and we had to remove them from the dataset. In addition, more complex social interactions may lead to the engagement of additional cognitive processes that vary from person to person due to using different strategies to make their lies more convincing. For example, in the follow-up question of what strategies they used to tell convincing lies, some participants reported they focused on facial expression, some focused on the tone of voice, and yet others focused on answering speed.

It is noteworthy to mention that several participants tried to gain some trust from the judge by telling the truth at the beginning of the game, as reported in the follow-up questions. This strategy has been considered to be one of the key components of deception proposed as "impression management" (Sip et al., 2008). To deceive the other successfully, some foundation in truth is required so that people tend to incorporate deception with truthful responses instead of making a series of exclusively deceptive responses (DePaulo & Kashy, 1998; DePaulo et al., 1996; Ekman, 2009). Establishing a reputation for being trustworthy by building the initial cooperation is necessary for deception in order to secure the trust of receivers (Sip et al., 2012). Based on the strategy used by participants, the current design successfully engaged a series of cognitive operations required for participants to conduct deceptive communication in a social context.

Detection of deception

In addition to the stimulus and judgment type, the individual differences in detecting lies based on their behavioral performance (i.e., Good judge vs. Bad judge) were also studied. The results showed that there was no electrocortical signature for stimulus and judgment types across the groups. In other words, participants' ERPs revealed no difference between whether the person in the video was trying to deceive or to tell the truth, or between whether participants judged the person as truthful or deceptive. Similarly, there was no difference between good and bad judges in any comparisons of stimulus and judgment types.

It has been reported repeatedly that detection of deception using behavioral cues is difficult (DePaulo et al., 1980; Ekman & O'sullivan, 1991; Vrij, 2000) with people's ability to detect lies being just slightly better than chance (i.e., 54% on average) (Bond & DePaulo, 2006). Surprisingly, only a few previous studies have examined the neural correlates of detecting deception. One study was conducted by Sip et al. (2010) as mentioned earlier, in which participants had to judge whether the opponent was being truthful or not in the Meyer game. They found no differences in neural activity between judging opponent as truthful versus deceptive regardless of whether the judgment was correct or not. Similarly, our study found no difference between judgment types across groups (i.e., good vs. bad judges).

In an fMRI study conducted by Grèzes et al. (2004), participants attempted to detect whether an actor in the clip was pretending that a box they lifted was heavier or lighter than it actually was. They observed increased activity in the amygdala when participants inferred deceptive intention from the actions of the actor regardless of whether the judgment was correct or not, possibly reflecting an emotional response to the belief of being deceived. They also observed increased activity in right superior temporal sulcus (STS), an area involved in the perception of biological motion (Allison, Puce, & McCarthy, 2000). However, our study found no ERP differences between judgment types. One possible reason for this is that the current study participants could only judge the person in the video based on their vocal answer with perhaps some minor facial expressions rather than their full-body behavior. As in Grèzes et al. (2004) study, the activity in STS might reflect inferential judgment about a person's intention based on observed body movements. Another possible reason is that neural activity elicited by being deceived triggered emotional responses like Grèzes and colleagues have found (i.e., an enhanced activity in the amygdala), but that this subcortical activity was too subtle for the ERP technique to detect. It is also possible that the detection of deception task in the current study did not prompt many emotional responses; instead, some participants might focus on the evaluation of probabilities of the card the person reported. For example, if the individual reported the high-value card much more frequently than the other cards, participants would become aware of it and take it into consideration for their judgment. In other words, for some participants, the strategies for detecting lies may be based more on perceived probabilities of the card value rather than simulating the intention of the person.

Comparison with the head movement task

The comparison between the detection and head movement tasks illustrated the ERP differences between higher-level and lower-level cognitive processes. The results showed that the head movement task elicited a larger N1 than detection task. The visual N1 component is typically modulated by visual attention processes (Luck & Yard, 1995; Vogel & Luck, 2000). Paying attention to nonspatial features such as colour, motion or shape modulates occipital-parietal ERPs between 100 and 150 ms (Hillyard & Anllo-Vento, 1998). Wang, Kuroiwa, Li, Wang, and Kamitani (2001) also suggested that the N1 might be involved in both early and late visuo-spatial processing, whereas the preceding P1 might reflect early perceptual processing. Thus, the larger N1 in the head movement than the detection task may reflect visual attention to head motion.

The lie detection task required participants to observe the behavioural cues from the person in the clip, and the behavioural cues participants relied on in making their judgment varied. Ekman and O'sullivan (1991) gave participants open-ended descriptions of behavioural clues they used in judging whether the person in the video was lying or not. Many answers were given, including answering too slowly, strained voice, avoiding eye contact, phony smile and body language. Thus, unlike the head movement task with a clear and uniform target to attend, in the lie detection task participants could use different strategies to make the judgement according to their previous experience, thus affecting different ERP components when they were detecting deception. In our data, for the clip onset, the difference in mean amplitude between the two tasks appeared early at posterior sites and later shifted toward anterior sites. The initially difference at posterior sites was contributed by visual ERP components, whereas the subsequently difference at anterior sites might reflect higher-level cognitive activity involved in detecting deceit compared with the simple detection of physical movement

Comparison with the results in Study 1

The same SDT analyses used in Study 1 were applied in the current study to examine the individual differences in the ability to generate lies and to detect lies. No correlation was found between the ability to generate lies and the ability to detect lies ($r = -.12$, $p = .559$) as found in Study 1 ($r = -.099$, $p = .345$). Both results go against the idea of a deception-general ability proposed by Wright et al. (2012). Although the paradigm used in the current ERP study was completely different from the one in Wright et al. (2012), the existence of a deception-general ability should be tested in another context to show whether this phenomenon is independent of the paradigm used. To some extent, the current result provides such a test, further suggesting that these two abilities are not correlated, and providing no support for the idea of a deception-general ability.

Study 1 demonstrated that the ability to generate lies was higher in people with higher creativity achievement scores and who generated more uncued lies in the task. In order to examine whether this also was true in the current study, we correlated d'_{Sender} /$d'_{Receiver}$ with creativity scores and number of lies generated during the task, but neither d'_{Sender} nor $d'_{Receiver}$ correlated with creativity scores nor number of lies. There are a number of potential reasons for this difference. First, the hypothesis of the link between the ability to generate lies and creativity relies on the idea that both are involved in breaking standard rules to build associations between previously unassociated cognitive elements. We expected that people with higher creativity would create novel associations quickly (i.e., false information) and make their lies more convincing. The paradigm in Study 1 provided ideal circumstances for participants to lie because they were given 20 seconds to justify their opinions or describe their previous events and feelings. In contrast, the current paradigm only allowed the participants to speak one word in a physically constrained situation (i.e., EEG cap and electrodes) where there was no room for them to create unusual mental associations for the purpose of lying. Second, in order to make more convincing lies, instead of creating novel associations, the current paradigm was more likely to require the mental computation of reward maximization (i.e., monetary reward) by allocating deceptive and truthful responses based on the card type. In other words, the deception elicited in the current paradigm relied more on the decision-makings based on reward computation, which is another important lying ability to evaluate whether deception is advantageous based on the current situation.

Limitations and future studies

This study successfully created a realistic experimental environment to investigate deception closer to real life than in previous paradigms. However, the ecologically valid paradigm comes at a price. The first limitation is that we could not control how many deceptive and truthful responses were generated by the participants. Thus, some participants who generated only few lies had to be excluded from the analysis. Also, individual differences in how frequently people told lies resulted in a broad range of total lies from 15 to 97, which increased the variability in our data. Another limitation is that we did not know the intention of participants when they made truthful responses. Based on the bluffing game rule, if a truthful response was rejected, participants would get compensation on top of the card value. It was possible that participants made truthful responses with deceptive intent and expression in order to mislead the judge and get compensation. In future studies, one could manipulate this effect by providing compensation or not in false alarm trials. Then one could examine if the intention of inducing misjudgement exists and how this intention influences the ERPs accordingly.

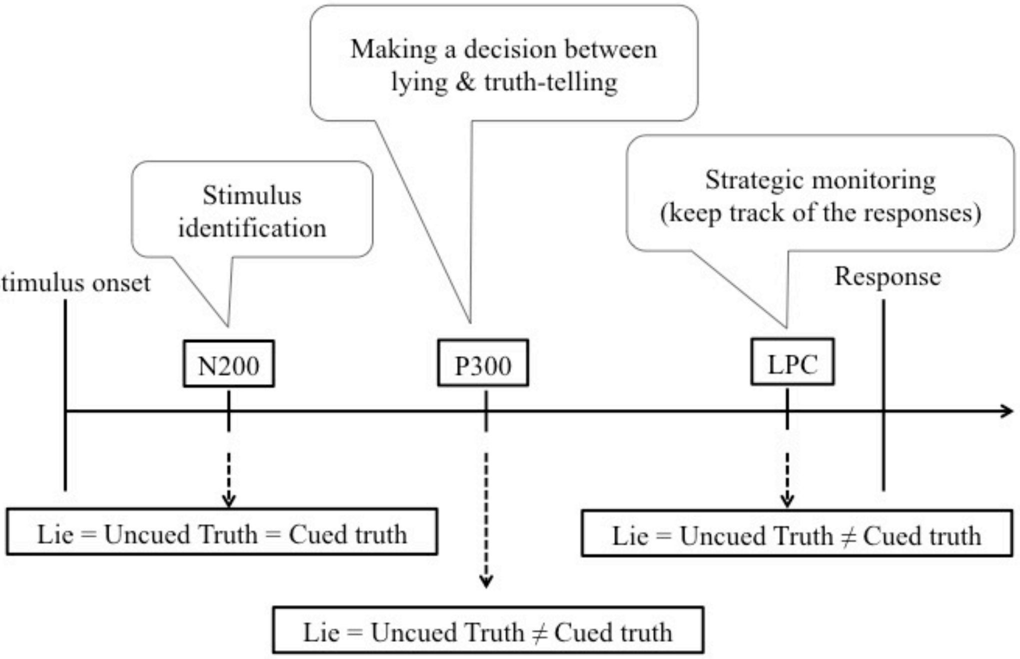
One possibility to improve the paradigm is for participants to play the game with a peer, not with an experimenter. According to the experimenter demand effects (Zizzo, 2010), the vertical relationship between the experimenter and participant can elicit different behaviours from a horizontal peer-to-peer interaction. Especially in social interactions, participants would more likely behave in a certain way to fulfil the expected objectives of the experiment or receive more pressure during the interaction. To minimize this effect, in future studies participants can be informed that they will play with a peer and experimenters have no access to their interaction until the end of the task. Although the interaction between participants and the experimenter (i.e., judge) was not face-to-face, avoiding this possible effect would increase the external validity of the experiment.

Additional limitations for the current lie detection task for investigating the neural processes underlying detection of deception are worth mentioning. First, each trial lasted around 1.5 seconds limiting nonverbal cues, and the verbal cue was only a single word from the person in the clips. Thus, instead of making judgements based on deceptive behavioural cues generated by the person, participants might focus on the context, such as the evaluation of probabilities mentioned earlier. Future research needs to create a set of stimuli with considerable verbal and nonverbal cues in order to explore the neural correlates underlying detection of deception. Second, some participants reported that they lost their attention at some point because they had no idea whether their judgment was correct or not. Some participants also reported that it was more difficult to concentrate on the lie detection task compared with the bluffing game since there was no monetary motivation for a correct judgment. In future studies, one could provide feedback on each trial. In this design, it is likely that participants would modify their strategies constantly according to the feedback so that intra-subject variability would increase. However, it would be interesting to examine whether participants' ability to detect lies would improve as a result of feedback.

Summary and Conclusion

The current study provided participants the choice to deceive by creating a context where deception was possible in some conditions but also came with the risk of being punished if it was detected. Our paradigm captured the idea that real-life deception is a decision with costs and benefits in which the deceptive agents have to decide between the possible reward from deception and the cost of being caught. The paradigm allowed us to investigate deception as the outcome of social decision-making. Our results suggested that deceptive behaviour commences with a target-recognition process followed by a decision-making process. That is, the deceptive agent has to recognize the target to see whether deception is a possible course of action and then make a choice between lying and truth- telling. These processes in the early stage require similar cognitive resources regardless of whether the outcome is a lie or the truth as indexed by identical N200 and P300 for lying and uncued truths. Similar RTs for lies and uncued truths also support this suggestion. Once the spontaneous decision has been made, it requires strategic monitoring to keep track of the responses in order to maximize the gains regardless of whether the outcome is a lie or the truth as indexed by identical LPC for lies and uncued truths. These processes are summarised in Figure 3-9 suggesting that spontaneous deception and spontaneous/uncued truths require comparable cognitive effort. The lie detection task was a first attempt to examine the electrophysiological signals underlying detection of deceit. Our result showed no ERP differences in any comparisons. Future studies are needed to investigate whether there are neural signatures of correct and incorrect judgments using stimuli with both verbal and nonverbal cues. Finally, in line with the results in Study 1, the ability to generate lies was not correlated with the ability to detect lies against the idea of a deception-general ability.

Figure 3-9. A diagram depicting some of the suggested cognitive processes for lying compared with two types of truth telling



CHAPTER 4

STUDY 3 – The Effect of Mental Countermeasures on ERP-based Concealed Information Test

Introduction

The last ERP study reported in Study 2 has elucidated key cognitive and neural processes underlying the generation of uncued lies by implementing an interactive card game where the deceptive agents had to evaluate possible reward and punishment and make a decision of whether to lie or not. The results have provided theoretical understanding of neural mechanisms underlying real-life deception. In the current ERP study, we focused on applied end by examining the validity of the lie detection paradigm, i.e., concealed information test, when applying mental countermeasures, and determining the role of creativity in countermeasure use.

Concealed information tests (CITs, also known as Guilty Knowledge Tests, were introduced by Lykken (1959, 1960) have been used for many decades to determine the presence of crime-related knowledge in a suspect's memory by measuring behavioural, autonomic, electrophysiological and hemodynamic variables (Ganis et al., 2016; Ganis et al., 2011; Meijer et al., 2014; Rosenfeld, Ben-Shakhar, & Ganis, 2012; Verschuere et al., 2011). The most common type of CIT paradigm employs 3 types of stimuli (Rosenfeld et al., 2012): probes (familiar items from a crime scene, such as a stolen ring), irrelevants (control, unfamiliar but comparable items, such as a necklace, a bracelet, a watch, etc.) and a target (an item with an unique response requirement for holding attention). The main comparison of interest in this 3-stimulus paradigm (3-S CIT) is between the probe and the irrelevants (normally, the mean of the irrelevants). The logic of this paradigm is that if the probe consistently elicits a stronger response than the irrelevants in a person, then one can infer that this person is familiar with the probe even though the person denies having knowledge of the probe: Deception may be taking place under this circumstance. Conversely, if there is no different response between probe and irrelevants, one may conclude that the person has no knowledge of the probe.

The detection of deception using CITs relies on solid scientific ground (e.g., Verschuere, Ben-Shakhar, and Meijer 2011), mostly from psychophysiological research on the orienting reflex (OR) (e.g., Siddle 1991, Sokolov 1966, Sokolov 1963). The OR is a complex of physiological and behavioral reactions evoked by significant stimuli (e.g., Gati and Ben-Shakhar 1990, Siddle 1991). Traditionally, the responses examined have been autonomic nervous system (ANS) measures, such as skin conductance (SCR), respiration and heart rate (for reviews, see Meijer et al., 2014, Ben-Shakhar and Elaad 2003). The typical responses elicited by the familiar CIT items in knowledgeable individuals are increased SCR (Lykken, 1959), heart rate deceleration (Verschuere, Crombez, De Clercq, & Koster, 2004) and respiratory suppression (Timm, 1982). Meta-analyses have shown that SCR measures can discriminate between individuals with and without knowledge with high accuracy (effect size (Cohen's *d*) of 1.55 for both Meijer et al., 2014 and Ben-Shakhar & Elaad, 2003).

In more recent years, event-related potential (ERP) have been used in CIT paradigms, the focus of this study (for reviews, see Rosenfeld 2011, Rosenfeld, Ben Shakhar, and Ganis 2012). Two main ERP components have been investigated in these paradigms, the anterior N2 and the P300.

With regard to the N2 (Gamer & Berti, 2010, 2012; Ganis & Schendan, 2012; Hu, Pornpattananangkul, & Rosenfeld, 2013), the findings so far have been inconsistent. Gamer and Berti (2010) reported a larger N2 to the probe, which was attributed to an orienting reflex to meaningful information. However, no such difference was found in a second study (Gamer & Berti 2012). Another study found the opposite effect, with a larger N2 to the irrelevants than the probe (Ganis & Schendan, 2012). Ganis et al., (2016) suggested that the overall N2 effect observed in CIT studies with visual stimuli may be the superposition of multiple N2 components with similar latency but modulated in opposite ways by different factors, such as stimulus complexity and the similarity between stimuli. Therefore, the N2 component is not a straightforward index of concealed information. In the current study, the N2 component will be measured in order to compare it with previous CIT studies but only the P300 component will be used for further classification analyses.

The P300 component has been mostly investigated as a robust index of recognition of concealed information, with the probe eliciting a P300 with a larger amplitude than the irrelevants (e.g., Meijer et al., 2014, Rosenfeld et al. 1988, Rosenfeld et al. 2004). The P300 is a positive brain potential, which usually occurs between 300 and 800 ms after stimulus onset. Typically, the P300 is elicited by rare stimuli in a series of standard stimuli in an oddball paradigm and its amplitude is inversely related to the subjective frequency of the eliciting stimulus (Donchin & Coles, 1988). Furthermore, its amplitude increases with the level of significance of the stimulus (Berlad & Pratt, 1995; Johnston, Miller, & Burleson, 1986). The P300 has two subcomponents: (i) the P3a, thought to reflect stimulus-driven frontal attention mechanisms and is characterized by a frontal/central scalp distribution; (ii) the P3b, thought to originate from temporal-parietal of attention and subsequent memory processing, and characterized by a parietal scalp distribution (Polich, 2007). The P300 described here is an instance of the P3b, and the term “P300” will be used throughout the chapter.

In the CIT paradigm, the P300 is enhanced in individuals with concealed knowledge by infrequently presenting the probe (e.g., $p = .17$) within a stream of frequent irrelevants (e.g., $p = .67$). However, to individuals without probe knowledge (innocent), the probe is just another irrelevant that elicits a smaller P300, comparable to that of other irrelevants. A meta-analysis (Meijer et al., 2014) examined the validity of the CIT using P300 amplitude and showed high accuracy in differentiating between individuals with and without probe knowledge (effect size (d^*)= 1.89). However, the P300 is not elicited by the recognition of concealed information, but by many other factors, for example, anything that affects the saliency of a stimulus (Polich, 2012). Thus, a large P300 to an item does not necessarily mean that one is familiar with that item. Such lack of selectivity means that P300 amplitude as an index of concealed knowledge is potential vulnerable to countermeasures. The main aim of this study was to determine the effect of cognitive countermeasures on P300 amplitude in a CIT paradigm.

To determine whether CIT paradigms can be useful in real-life settings, it is necessary to investigate the effect of countermeasures (CM), deliberate strategies used by suspects in order to alter their psychophysiological reactions during the test (Ben-Shakhar, 2011). Since the contrast between the probe (i.e., crime-related item) and the irrelevants (i.e., neutral items) is the key index to discriminate between individuals with concealed knowledge and no knowledge, countermeasures can be employed to suppress the responses to the probe or to artificially enhance responses to the irrelevants (cf. Ben-Shakhar, 2011). Suppressing autonomic responses to the probe is relatively difficult because such a strategy may increase item saliency and end up producing larger responses (Elaad & Ben-Shakhar, 1991). Thus, most studies on the topic employed the later CM strategy of trying to enhance responses to the irrelevants during the CIT. The CM techniques reported in the literature can be categorized into physical CMs (e.g., Rosenfeld et al., 2004 using covert physical acts, like pressing the left forefinger), and mental CM, (e.g., Sasaki, Hira, & Matsuda, 2001 counting backwards by sevens). Note that some CMs may be effective with some measures (e.g., skin conductance, as in Honts et al., 1996) but not with others (e.g., ERPs, as in Sasak et al, 2001). Mental CMs are generally more difficult to detect than physical ones because they do not result in behavioral differences one can easily measure. Thus, it is especially important to investigate to what extent the validity of the test is affected by the mental CMs since they are hardly detectable. The current study aimed at examing how mental CMs affect concealed knowledge detection using P300 amplitude.

Recent studies have begun to examine how CMs can affect the accuracy of a 3-S CIT paradigm using ERP and fMRI measures which, at first sight, may seem more resistant to CMs than ANS measures. Sasaki et al. (2001) instructed participants to count backward by sevens in a P300-based CIT and found this mental CM was largely ineffective based on P300 amplitude. Later on, Rosenfeld et al. (2004) devised a more effective CM strategy that involved making irrelevant stimuli meaningful by assigning covert responses to each irrelevant, including some physical (e.g., imperceptibly wiggling the big toe in the left shoe) and one mental CM (i.e., imaging the experimenter slapping the participant in the face). This combination of physical and mental CMs significantly decreased the hit rate from 82% (normal guilty group) to 18% (CM guilty group). This study showed that concealed knowledge detection using P300 amplitude is readily defeated by a mixture of physical and mental CMs. An fMRI study conducted by Ganis et al. (2011) employed physical CMs like those used by Rosenfeld and collaborators (2004) in a 3-stimulus CIT paradigm and found that hit rates of concealed knowledge detection fell to 33% with CM from 100% without CM using activation in ventrolateral and medial prefrontal cortices.

These studies have shown that physical CMs can significantly impair classification accuracy using P300-based or fMRI-based measures in 3S CIT paradigms. However, it remains unknown whether purely mental CMs can also decrease the accuracy of P300-based CITs. Although Rosenfeld et al. (2004) implemented one mental CM, the proportion of mental CMs (20%) was small compared with physical CMs (80%) and so the effect was likely driven by the physical CMs. This is an important question because mental CMs are unlikely to be detected by the examiner, even with sophisticated instrumentation (Elaad & Ben-Shakhar, 1991). The aim of this ERP study was to determine whether purely mental CMs are effective in 3S CITs. Mental CMs employed in the current study aimed at both suppressing responses to the crime-related item and enhancing responses to neutral items. By reducing the response difference between these two types of stimuli, we expected to reduce the difference in P300 amplitude between the probe and irrelevants, thus decreasing classification accuracy.

Another question we wanted to address was whether CM efficacy is linked to certain cognitive abilities. It has been found that individuals with higher standard creativity measures have shown to have more flexible cognitive control (Zabelina & Robinson, 2010). In the current study, participants were instructed to apply different mental coutnermeasures to the probe and the irrelevants respectively (for details, see Materials and methods). We hypothsized that more creative individuals may be better at implementing this type of mental countermeasures as they may be better at switching from trial to trial between the two attentional modes requried by the countermeasures.

Finally, note that variants of the 3-S CIT aimed at defeating saliency-based CMs have been under development (Bowman, Filetti, Alsufyani, Janssen, & Su, 2014; Winograd & Rosenfeld, 2011). However, given that these methods are quite new and have not yet been systematically replicated by multiple laboratories, they will not be discussed in this thesis.

Materials and methods

Participants

A total of 45 healthy participants (32 females and 13 males, mean age = 20.6 years, SD = 2.3), recruited from University of Plymouth, completed no knowledge (NK), concealed knowledge (CK) and countermeasure (CM) tasks. All participants had normal or corrected vision, and no history of neurological or psychiatric disease. Data from 2 participants were excluded from the analyses (one due to excessive muscle tension artifacts and the other due to misunderstanding of the rules), leaving 43 participants (30 females and 13 males, mean age = 20.7 years, SD = 2.3). All procedures were approved by the Plymouth University Faculty of Science and Technology Human Ethics Committee. All participants gave signed informed consent and were remunerated for their time in the study.

Stimuli

Stimuli were the single digits 3, 4, 5, 6, 7, and 8, green against a black background. Probes and targets varied across participants and these pairs were drawn randomly from the 20 possible pairings with a numerical distance between the probe and the target of at least 2. The same probe-target pair was used for the NK, CK and CM conditions. Each digit subtended approximately 2 x 2 degrees of visual angle (Figure 4-1).

Figure 4-1. Digit stimuli

Procedure

The study was divided into two sessions. The first session was the EEG experiment where participants completed the NK, CK and CM tasks. After the EEG experiment, participants were scheduled to come back for a second session to complete cognitive tests and questionnaires.

After setting up the EEG cap and electrodes, participants were seated on a comfortable chair, 115 cm from a computer screen in a dark room. The paradigm was presented by Presentation® software (Version 18.0, Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com). A total of 210 stimuli were presented in each condition with each digit presented 35 times. Each stimulus was presented for 800 ms, with an average inter trial interval (ITI) of 3500 ms, randomly varying between 3200 and 4000 ms to minimize expectation potentials.

At the beginning of the study, participants were given a target digit verbally and instructed to press a 'Yes' button to it with their dominant hand during the entire experiment. In the NK condition, participants pressed a 'Yes' button when they saw the target digit and a 'No' button for all other digits. Before starting the second task (CK condition), participants were instructed to choose a "secret" digit (probe) by selecting an envelope from a set of 5 (all containing the same digit assigned by design to that participant). The instructions asked participants to open the selected envelope, take the card with the secret digit inside it, keep it face down on their lap, and to lie by denying having knowledge of this digit by pressing the ‘No’ button throughout the test. Participants were informed that the experimenter did not know their secret digit (probe) as they kept the envelope until the end of the experiment. Their job was to lie to the experimenter by trying to control their behavior and brain signals, as these would be monitored during the test for the experimenter to find signs of deception. In the final task (CM condition), participants were instructed to perform a mental countermeasure composed of two strategies to beat the test while performing the CK task just described. The first strategy was to pay attention to superficial aspects of the probe by focusing on some physical features (e.g., font, color or size) of the digit. In other words, participants were asked to try to ignore the meaning of the probe and to focus on how it looked like on the screen. The second strategy was to think of something meaningful from one’s memory that was associated with each of the irrelevant digits and once the association between one’s own memory and each irrelevant digit was created, to just think of the same memory each time one encountered the corresponding irrelevant. This strategy was designed to try to make the irrelevant digits more meaningful to the participants thus reducing the difference between the secret digit and the other irrelevant digits.

To ensure that the secret number had no meaning in the NK condition and no strategies were applied in the CK condition, all participants completed the three tasks in the same order, i.e., NK, CK and CM. Prior to each task, participants carried out a practice session to make sure that they understood the instructions and performed the tasks correctly. The same stimulus sequence was used in the three tasks for the practice session and it consisted of 12 trials.

In order to avoid confounds due to local statistical properties of the sequences, the stimulus sequences used in these three tasks were the same. Thus, the pattern of button presses was identical in each task and the only differences between tasks were the concealed information about the probe and the execution of the CMs.

Creativity tests

Participants came back for the second session individually to complete the Wechsler Abbreviated Scale of Intelligence (Wechsler, 1999) and two creativity tests, i.e., the Abbreviate Torrance Test for Adults (ATTA) (Goff, 2002) and the Creativity Achievement Questionnaire (CAQ) (Carson et al., 2005).

Electrophysiological data acquisition

The electroencephalogram (EEG) was sampled at 8192 Hz using a Biosemi ActiveTwo system. EEG data were collected from 32 electrodes Ag/AgCl electrodes arranged according to the 10-20 system, and loose lead electrodes below the left eye to monitor eye blinks and vertical eye movements, and on the left and right mastoids. Horizontal eye movements were monitored using 2 loose electrodes places on the outer canthi of the right and left eyes. The data were downsampled off-line to 512 Hz before further processing. For analyses, data were re-referenced off-line to the average of the two mastoids for consistency with most of the existing literature on the topic.

ERP analyses

ERPs were averaged off-line for an epoch of 1200 ms (including a 200 ms baseline), beginning 200 ms prior to stimulus onset. Independent component analysis (ICA) was conducted for 65% of participants using the *runica* algorithm. Independent component analysis-based artefact identification methods were employed to identify ocular and muscle artifacts (Mognon et al., 2011; Winkler et al., 2011). The results of the artefact detection were visually inspected before the components were manually removed. Subsequently, residual trials contaminated by blinks, eye movement, muscle activity or amplifier blocking were rejected off-line. The data were low-pass filtered at

30 Hz.

Analyses focused on the comparison between the probe and the irrelevants. Here we use the term “probe effect” to indicate the difference in the measure of interest between the probe and irrelevants. The P300 probe effect, regarded as a robust index of recognition, was used here to determine whether and to what extent the mental countermeasures worked. For completeness, we also measured the amplitude of the N2. The time window used to quantify P300 amplitude was between 500 and 700 ms. The amplitude of the N2 component was measured between 200 and 250 ms. The selection of time windows was based on the results of previous research (e.g., Farwell & Donchin, 1991; Ganis et al., 2016; Rosenfeld, 2011; Rosenfeld et al., 2004) as well as on aggregated grand average from trials.

To assess the overall pattern of results, repeated-measures ANOVAs on lateral and midline sites were conducted on the mean amplitudes of the average ERPs from each participant. The “lateral” ANOVAs were carried out on the 14 pairs of lateral sites using three factors: Item Type (probe vs. irrelevants), Hemisphere (L vs. R) and Site (14 site pairs). The “midline” ANOVAs were carried out on the 4 midline sites (Fz, Cz, Pz, and Oz) using 2 factors: Item type (probe vs. irrelevants) and Site (4 midline sites). The “irrelevants” level in the item type factor was the average of all 4 irrelevants.

After the omnibus analyses, we focused on the Pz site for further analyses of the P300, as Pz has been repeatedly reported in the previous studies (Ganis et al., 2016; Rosenfeld, 2011; Rosenfeld et al., 2004) and also showed the largest effect of concealed information in the current study compared with other sites. To assess if countermeasures reduced the P300 probe effect, we directly compared the activity of (Probe – Irrelevants) between CK and CM condition by conducting paired t-tests. To examine if the reduced P300 probe effect was due to a decreased P300 to the probe or to an increased P300 to irrelevants, repeated measures ANOVAs were conducted using two factors: Condition (CK vs. CM) and Item Type (probe vs. irrelevants). Follow-up t-tests were carried out to compare pairs of items of interest.

Receiver Operating Characteristics (ROC) and linear discriminant analyses were conducted between each group (NK vs. CK, NK vs. CM and CK vs. CM) to determine the accuracy with which the P300 component could be used to classify between different conditions. We also compared the area under the ROC curve (AUC) between for the classification of NK vs. CK and NK vs. CM using a nonparametric approach (Hanley & Hajian-Tilaki, 1997). These ROC analyses were performed using MedCalc, version 12.0 (MedCalc Software, Ostend, Belgium).

Correlations between ERP component and creativity measures

Since we hypothesized that more creative individuals might implement countermeasures more efficiently because of higher flexible cognitive control, we conducted correlation analyses between the mean amplitude of the P300 probe effect and creativity measures. We expected to see individuals with higher creativity score showing smaller P300 probe effect in the CM task.

Results

Behavioral results

For the response time (RT), a two-way ANOVA was conducted using 2 factors: Item type (probe, irrelevants) and Condition (CK, CM, NK). The results showed a main effect of item type, $F(1, 42) = 39.62$, $p < 0.001$, $\eta^2 = 0.485$, with participants responding slower to the probe than the irrelevants, and a main effect of condition, $F(2, 84) = 12.53$, $p < 0.001$, $\eta^2 = 0.230$, with participants responding slower in the CM compared with CK and NK condition but no difference between CK and NK. There was also a significant interaction between item type and condition, $F(2, 84) = 16.11$, $p < 0.001$, $\eta^2 = 0.277$, indicating that the effect of item type (i.e., probe and irrelevants) differed across conditions. Although the pattern of YES/NO responses was constant across conditions, the RTs revealed that participants followed the instructions of each condition and performed differently in the different conditions.

Next, to understand how the effect of item type differed across conditions, we conducted paired t-tests to compare RTs between probe and irrelevants in each condition (see Table 4-1 for descriptive statistics). The RTs showed no significant difference between the probe and irrelevants in the NK condition, $t(42) = -1.052$, $p = 0.299$. In contrast, the RTs were slower for the probe than the irrelevants in the CK condition, $t(42) = 5.431$, $p < 0.001$, $d = 0.83$. Thus, hiding information of probes in the CK condition required deceptive response that took more time than processing irrelevants (i.e., honest response). Similarly, the RTs showed slower responses for the probe than the irrelevants in the CM condition, $t(42) = 5.051$, $p < 0.001$, $d = 0.77$.

We also conducted paired t-tests to compare the RTs between CK and CM condition for the probes and irrelevants, respectively in order to examine the effect of countermeasures on the RTs. The RTs for the probe were slower in the CM than CK condition, $t(42) = -3.268$, $p < 0.005$, $d = 0.50$. Similarly, the RTs for irrelevants were slower in the CM than CK condition, $t(42) = -3.747$, $p < 0.001$, $d = 0.57$. The longer RTs in the CM condition suggest that participants were performing the countermeasures to make irrelevants more salient for them and the probe less significant by focusing on physical features.

Table 4-1

Mean (SD) response time (ms) in the NK, CK and CM conditions

	NK	CK	CM
Probe	455.7 (±65.9)	489.0 (±90.9)	564.9 (±178.6)
Irrelevants	459.8 (±68.9)	445.9 (±72.3)	506.2 (±129.0)

P300 (500-700 ms): No knowledge (NK) condition

Without probe knowledge, there was no P300 probe effect. Both lateral and midline ANOVAs revealed neither a main effect of item type nor any interactions involving item type (Table 4-2, and Figure 4-2a).

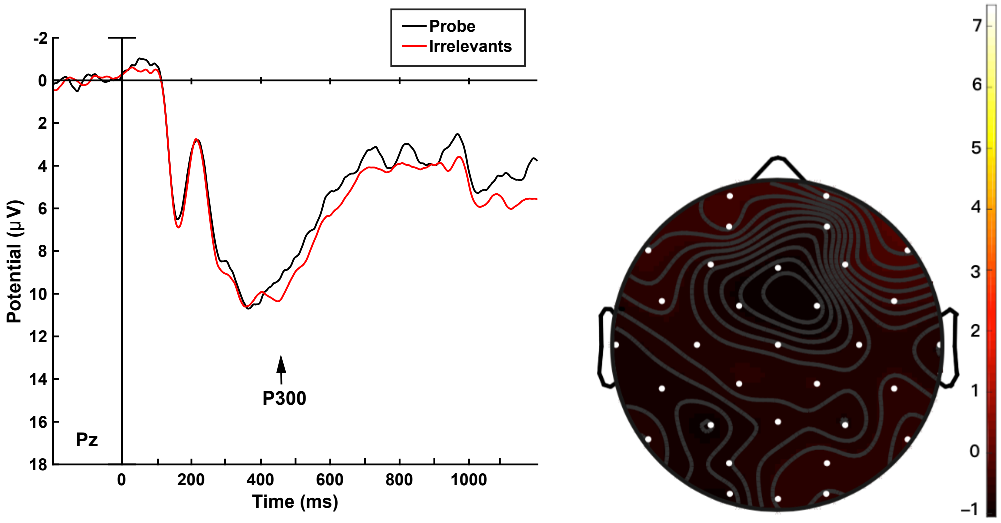
P300 (500-700 ms): Concealed knowledge (CK) condition

The standard CIT effect on the P300 (i.e., P300 probe effect) was replicated when participants were informed about a probe. Lateral and midline ANOVAs revealed both a main effect of item type and an interaction between item type and site (Table 4-2) where the probe was more positive than the irrelevants and the effect was maximal at the central-parietal sites. For the midline ANOVA, there was a main effect of item type, and an interaction between item type and site. The direct comparison between probe and irrelevants at each site (i.e., Fz, Cz, Pz and Oz) showed significant differences, $t(42) > 4$, $p < 0.001$ for all contrasts (Figure 4-2b shows the data at Pz as a representative site).

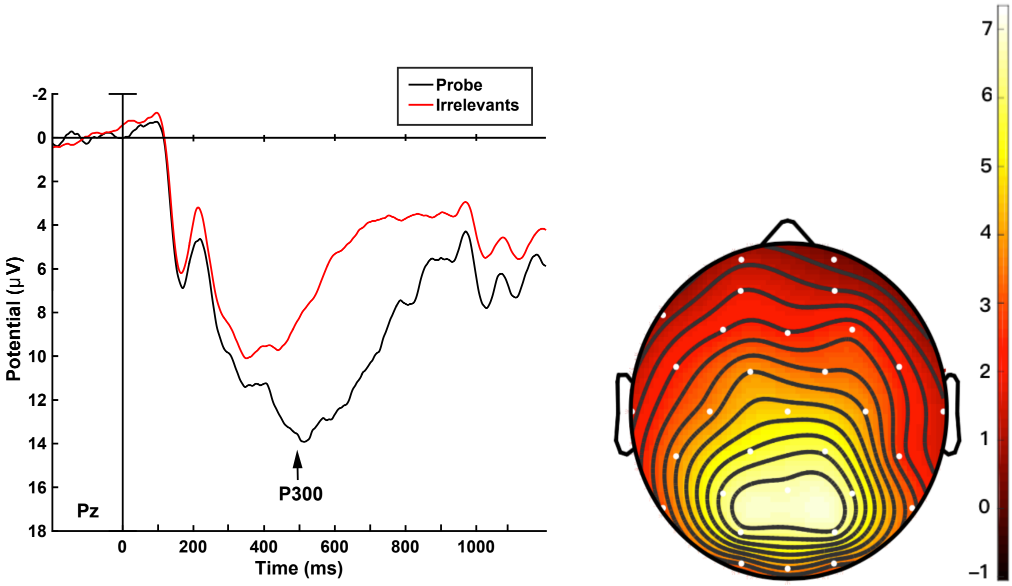
P300 (500-700 ms): Countermeasure (CM) condition

Although participants were instructed to implement countermeasures in this condition, the CIT effect on the P300 was still present. Lateral and midline ANOVAs revealed both a main effect of item type and an interaction between item type and site (Table 4-2) where the probe was more positive than the irrelevants, and the effect was maximal at the central-parietal sites. For the midline ANOVA, there was a main effect of item, and an interaction between item type and site. The direct comparison between probe and irrelevants at each site (i.e., Fz, Cz, Pz and Oz) showed significant differences, $t(42) > 2$, $p < 0.05$ for all contrasts (Figure 4-2c shows Pz site as a representative).

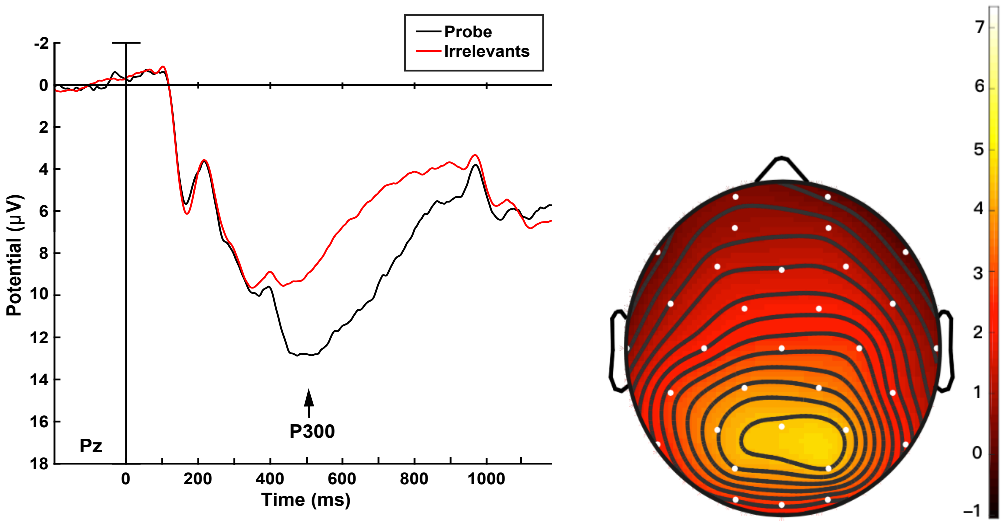
Figure 4-2. Grand-average ERPs for the probe and the irrelevants at Pz as well as the scalp distribution maps (500, 700 ms) for the P300 probe effect in the (a) NK, (b) CK and (c) CM conditions.


Probe
Irrelevants
-2
0
2
4
6
8
10
12
14
16
18
Potential (μV)
Pz
P300
0
200
400
600
800
1000
Time (ms)
7
6
5
4
3
2
1
0
−1

(a)


Probe
Irrelevants
-2
0
2
4
6
8
10
12
14
16
18
Potential (μV)
Pz
P300
0
200
400
600
800
1000
Time (ms)
7
6
5
4
3
2
1
0
−1

(b)


Probe
Irrelevants
-2
0
2
4
6
8
10
12
14
16
18
Potential (μV)
Pz
P300
0
200
400
600
800
1000
Time (ms)
7
6
5
4
3
2
1
0
−1

(c)

Direct comparison between Concealed knowledge (CK) and Countermeasure (CM) condition

We conducted an ANOVA using two factors (condition and item type) to examine whether the effect of item type differed between the CK and CM condition. There was a main effect of item type, $F(1,42) = 73.67$, $\eta_p^2 = 0.637$, $p < 0.001$, but no main effect of condition, $F(1,42) = 0.06$, $\eta_p^2 = 0.001$, $p = 0.806$. There was an interaction between condition and item type, $F(1,42) = 8.03$, $\eta_p^2 = 0.160$, $p < 0.01$, where the probe effect was larger in the CK than CM condition (Figure 4-3). Although the direct comparison between the CK and CM conditions for each item type separately showed no significant difference (probe: $t(42) = 1.22$, $p = 0.229$; irrelevants: $t(42) = -1.85$, $p = 0.072$) (Figure 4-4), the difference for the irrelevants approached significance.

In sum, the implementation of the countermeasure successfully reduced the P300 probe effect in the CM condition, and it was due to both a numerical reduction in P300 amplitude for the probe and an increase for the irrelevants.

Figure 4-3. Grand-average ERPs for the difference between the probe and irrelevants at Pz as well as the P300 scalp distribution maps (500, 700 ms)

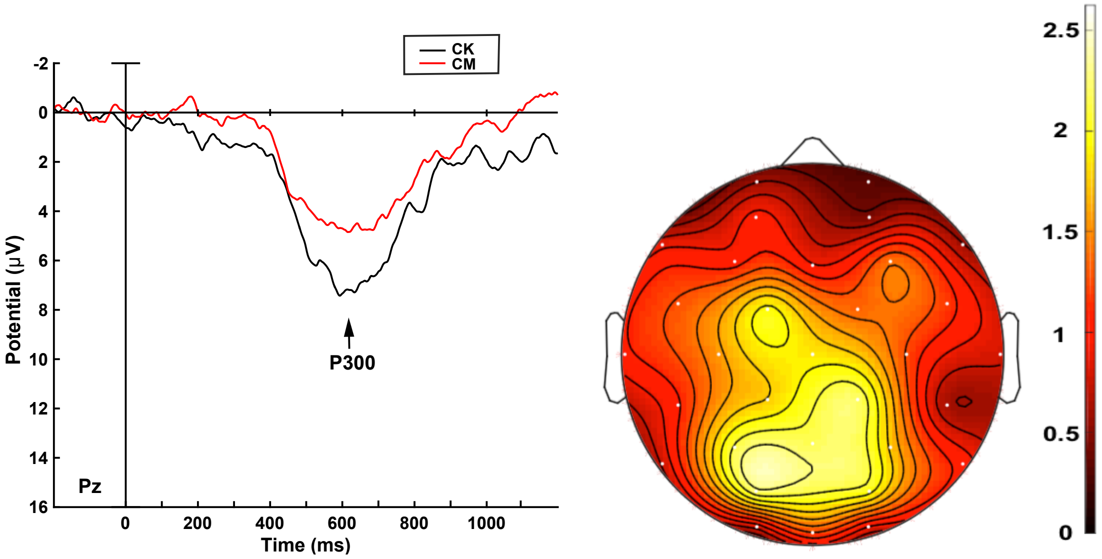
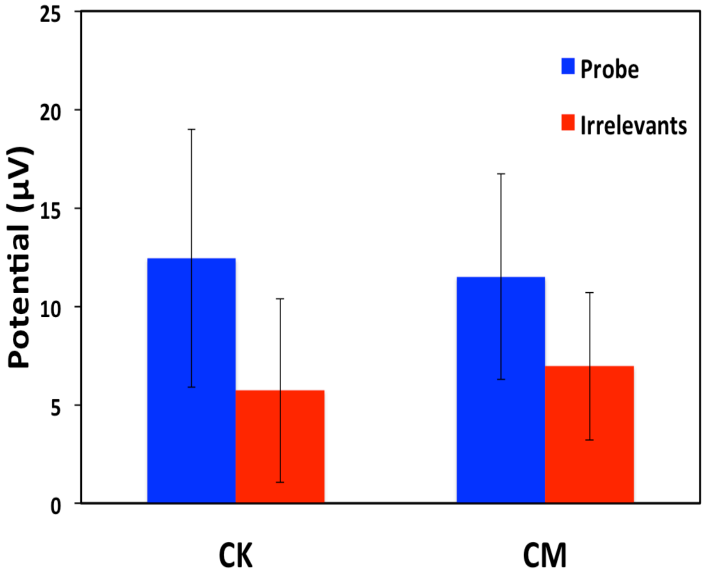


Figure 4-4. Mean potential (μV) for the probe and irrelevants in the CK and CM conditions at Pz



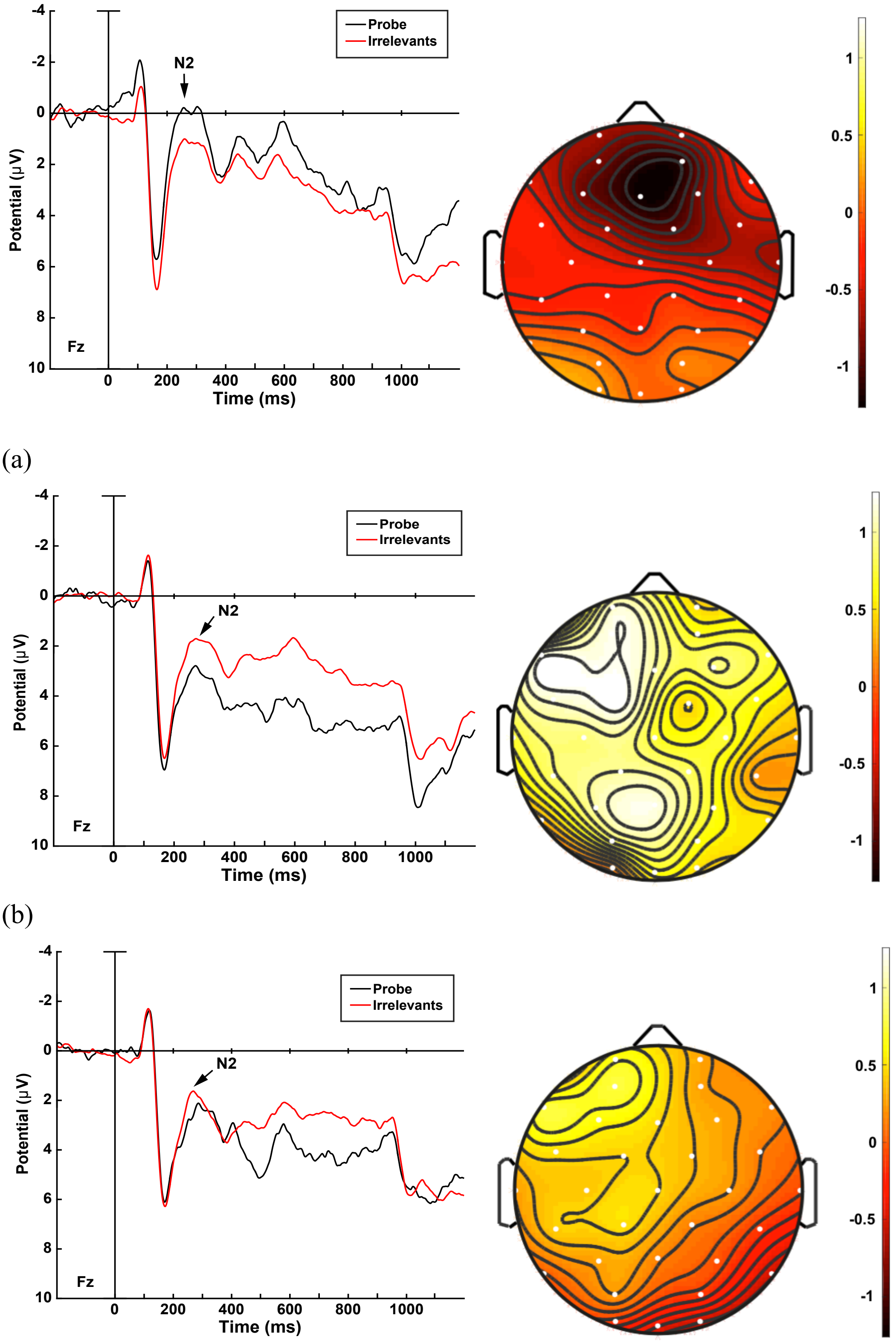
Classification between conditions

ROC analyses were carried out on the P300 probe effect at Pz between conditions. The area under the Curve (AUC) was 0.91 ([0.83, 0.96]) for the classification of NK and CK cases, 0.86 ([0.77, 0.93]) for the classification of NK and CM cases, and 0.63 ([0.51, 0.75]) for the classification of CK and CM cases. The AUC reduction due to CM use, relative to CK, was not statistically significant, but there was a trend towards significance ($p = 0.108$).

N2 (200-250 ms)

In the CK condition, both lateral and midline ANOVAs revealed a main effect of item type (Table 4-2) where the probe showed a smaller N2 than irrelevants (Figure 4-5b). This finding was in the same direction as in the study by Ganis and Schendan (2012) showing that concealed information is not necessarily associated with a larger frontal N2. In the NK condition, both lateral and midline ANOVAs revealed an interaction between item type and site (Table 4-2), indicating that the effect of item type varied from site to site (Figure 4-5a). In the CM condition, no main effect or interaction including the item factor were found (Figure 4-5c).

Figure 4-5. Grand-average ERPs for the probe and the irrelevants at Fz as well as the N2 scalp distribution maps (250, 300 ms) in the (a) NK, (b) CK, and (c) CM conditions.



(a)

(b)

(c)

Correlations between ERP component and creativity measures

To examine if more creative individuals implemented countermeasures more effectively, we conducted correlation analyses between the P300 probe effect at Pz site in the CM condition and creativity scores (i.e., CAQ arts, CAQ science and ATTA). No significant correlations were found. Only a trend towards significance in the expected direction was found with CAQ science ($r = -0.266$, $p = 0.084$).

Table 4-2

Lateral (Lat) and Midline (Mid) ANOVAs on N2 (200, 250 ms) and P300 (500, 700 ms) amplitude in no knowledge (NK), concealed knowledge (CK), countermeasure (CM) tasks, comparison between the CK and CM tasks and comparison between the CK and NK tasks. The item factor includes two levels: the probe and the average of the 4 irrelevants.

ERP	N2						P300					
	Lat			Mid			Lat			Mid		
Source	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2
No Knowledge (NK) Task												
Item	2.54	.12	.06	2.62	.11	.06	2.75	.11	.06	3.66	.06	.08
I x Site	**3.15**	<.05	.07	**5.71**	<.01	.12	0.84	.51	.02	0.36	.71	.01
I x Hemi	0.42	.52	.01				0.37	.55	.01			
Concealed Knowledge (CK) Task												
I	**8.08**	<.01	.16	**8.78**	<.01	.17	**55.70**	<.001	.57	**68.74**	<.001	.62
I x S	1.00	.38	.02	0.76	.46	.02	**27.29**	<.001	.39	**17.84**	<.001	.30
I x H	2.71	.11	.06				0.81	.37	.02			
Countermeasure (CM) Task												
I	0.57	.45	.01	0.27	.60	.01	**18.29**	<.001	.30	**26.42**	<.001	.39
I x S	1.67	.18	.04	2.42	.10	.06	**27.69**	<.001	.40	**14.55**	<.001	.26
I x H	**6.00**	<.05	.13				0.00	.95	.00			
Concealed Knowledge Task vs Countermeasure Task												
I	**5.23**	<.05	.11	**4.52**	<.05	.10	**49.65**	<.001	.54	**64.24**	<.001	.61
I x S	1.85	.16	.04	1.66	.20	.04	**39.79**	<.001	.49	**23.48**	<.001	.36
I x Task	1.23	.30	.03	2.35	.13	.05	**5.68**	<.05	.12	**6.65**	<.05	.14
I x H	**6.50**	<.05	.13				0.31	.58	.01			
I x S x H	1.01	.43	.02				**2.08**	<.05	.05			

Concealed Knowledge Task vs No Knowledge Task												
I	1.05	.31	.02	1.08	.31	.03	**15.00**	<.001	.26	**22.40**	<.001	.35
Task	**32.60**	<.001	.44	**28.74**	<.001	.41	**17.53**	<.001	.29	**24.79**	<.001	.37
I x S	0.86	.44	.02	1.80	.18	.04	**17.18**	<.001	.29	**15.92**	<.001	.28
I x T	**12.22**	<.001	.23	**12.48**	=.001	.23	**59.96**	<.001	.59	**84.67**	<.001	.67
I x S x T	**3.67**	<.05	.08	**4.69**	<.01	.10	**23.49**	<.001	.36	**11.08**	<.001	.21

Note. Degrees of freedom: i) No Knowledge, Concealed Knowledge, and Countermeasure Tasks. Item: 1, 42; Item x Hemi: 1, 42; Item x Site (Lat) & Item x Site x Hemi: 13, 546; Item x Site (Mid): 3, 126. ii) Concealed Knowledge Task vs Countermeasure Task & Concealed Knowledge Task vs No Knowledge Task. Item, Item x Hemisphere, Item x Task: 1, 42; Item x Site (Lat), Item x Site x Task (Lat), Item x Site x Hemisphere, Item x Site x Hemisphere x Task: 13, 546; Item x Site (Mid), Item x Site x Task (Mid): 3, 126. F values are with Greenhouse-Geisser correction.

Discussion

This study showed that the P300 probe effect could differentiate concealed knowledge and no-knowledge cases but that it became significantly smaller when participants applied mental countermeasures.

This result is novel and important because it is the first ERP study to apply purely mental CMs during a standard CIT. Previous ERP studies did not use CMs on the probe under the assumption that doing so would result in increased attention to the probe (Waid, Orne, Cook, & Orne, 1978; Waid, Orne, & Orne, 1981). Indeed, Elaad and

Ben-Shakhar (1991) found that item-specific countermeasures tended to increase psychophysiological detection with autonomic measures when participants were requested to count silently from one to ten every time the probe was presented. They proposed that this item-specific countermeasure requires participants to identify the probe prior to implementing the mental countermeasure and resulting in an increased autonomic response. In the current study, one strategy of the CM was applied to the probe, and it resulted in a numerical reduction of P300 amplitude compared to the CK condition, not an increase (Figure 4-4). This suggests that ERPs to the probe can be modulated via top-down inhibiting processes and that the probe is also vulnerable to mental countermeasures. The second CM strategy was applied to the irrelevants and it enhanced their saliency.

Compared with physical CMs, mental CMs are much less likely to be detected by examiners. Thus, it is important to examine whether ERP-based measure can still detect concealed information when the mental CMs are implemented. In this study, purely mental CMs appear to be less effective than the physical CMs (or a mixture of physical and mental CMs) used in previous studies because the AUC reduction only showed a trend towards significance. Previous CM studies using neuroimaging techniques have shown that combining physical and mental CMs, or using physical CMs only, significantly reduced concealed knowledge detection accuracy rate (Ganis et al., 2011; Rosenfeld et al., 2004). Although paired t-test showed the mental CMs successfully reduced the P300 probe effect in the current study, the classification did not show significance. Recently, Peth, Suchotzki, and Gamer (2016) systematically compared the application of physical and mental CM on autonomic measures during the CIT, and they found that physical CMs (e.g., moving their toes) were more effective than mental CM (e.g., imaging an emotional picture following specific irrelevant items). On reason for this difference may be that the effectiveness of mental CMs requires participants' imagination and the level of vividness may vary from trial to trial. This may enhance the inter- and intra- variability so that mental CMs only have a moderate effect on average. In contrast, physical CMs require participants to carry out specific movements of body parts when the corresponding items are shown. This type of CM is more likely to have a consistent effect on the dependent measures than mental CMs, as the movement is objective.

In a real crime testing, suspects will not know the irrelevant items until they start the test, as they did in the current study. Under this circumstance, it is unlikely that participants had practiced countermeasures beforehand. In one fMRI study conducted by Ganis et al. (2011), they instructed participants to imperceptibly wiggle the left index finger upon seeing the first irrelevant item, wiggle left the middle finger upon seeing the second irrelevant item, and wiggle the left big toe upon seeing the third irrelevant item. These consistent associations between covert actions and irrelevant items required some practice before participants executed them correctly. However, the suspects in a real crime case do not have the chance to practice which covert action is performed in response to each irrelevant item. What suspects could do is make spontaneous responses to the stimuli, just as participants did in the current study, where they had to generate mental associations upon seeing the irrelevant items. This strategy is more difficult to implement and is closer to a real-life situation, but it still affects the deception index (i.e., P300 amplitude) significantly, according to the current results.

For the anterior N2 results, we found that the probe elicited a smaller N2 than the irrelevants in the concealed knowledge condition, which is similar to what Ganis and Schendan (2012) found. This can be explained with the idea that the amplitude of the anterior N2 is proportional to the degree of mismatch to memory (Folstein & Van Petten, 2008; Folstein, Van Petten, & Rose, 2008). In the CK condition, compared with the probe, participants had minimal memories about irrelevant items, and so the N2 was larger to the irrelevants.

There was a small but significant interaction between item type and site in the

NK condition where participants did not have concealed information. One likely reason for this is that despite stimulus randomization across participants, residual perceptual differences between the probe and the irrelevants remained at the group level. This would be consistent with the idea that small perceptual differences among the stimuli can strongly modulate the N2 effect in the CIT paradigm, as shown in Ganis et al. (2016). These findings further confirm that the anterior N2 is not a robust index of concealed knowledge because it is affected by many cognitive processes, such as cognitive control, attention, novelty, etc. (Folstein & Van Petten, 2008).

The effect of mental CMs, as indexed by the P300 probe effect, was not associated with creative abilities, with the exception of a trend in the expected direction for the CAQ Science. In other words, the current result did not support the hypothesis that people with higher creativity scores were better at switching countermeasures from trial to trial between the two attentional modes, i.e., probe and irrelevants, respectively. It is possible that the ERP components involved in attentional switch occurred in the earlier time window, not in the P300 we focused on. From the previous ERP studies on visual selective attention, the P1 and N1 are frequently of interest and normally occur between 80 and 200 ms after stimulus onset (e.g., Hillyard & Anllo-Vento, 1998; Vogel & Luck, 2000). The P300 probe effect we examined occurred between 500 and 700 ms and its amplitude is mainly based on the perceived category probabilities of stimuli (see Johnson, 1988 for reviews), not attention orienting. Thus, due to the limitation of the dependent variable used in this ERP study, we were unable to find the smaller probe effect modulated by creative abilities. In the next study, we expected that the fMRI method would provide more information, e.g., a neural index representing the probe effect modulated by creativity, because of its better spatial resolution.

This is the first study to show how purely mental CMs influence the accuracy of P300-based CITs by reducing the difference between the probe and irrelevant. Although mental CMs seem to have limited effect compared with physical CMs, this study provides evidence that even with short response times (i.e., 800 ms), top-down cognitive processes can still affect the reliability of ERP-based CIT. This study also provides another general point that the vulnerability of the neuroimaging methods for deception detection to numerous countermeasures should be assessed, especially those countermeasures which cannot be observed by the examiners. The next study will continue examining the vulnerability of another neuroimaging method — functional magnetic resonance imaging (fMRI), for concealed knowledge detection using the same mental countermeasures.

CHAPTER 5

STUDY 4 – The Effect of Mental Countermeasures on fMRI-based Concealed Information Test

Introduction

This study used the same countermeasures employed in the previous chapter, but with an fMRI adaptation of the concealed information paradigm. The role of creativity in implementing these countermeasures was also examined.

Concealed information tests (CITs, also known as Guilty Knowledge Tests, invented by Lykken (1959, 1960) have been used for many decades to determine the presence or absence of crime-related knowledge in a suspect's memory (Ganis et al., 2011; Meijer et al., 2014; Rosenfeld et al., 2012; Verschuere et al., 2011). The basic logic of these paradigms (Lykken, 1959) is that recognition of an item of interest (referred to as "probe") will generate a stronger response (compared to suitable control items referred to as "irrelevants") that can be measured by monitoring behavioural, psychophysiological, or neural variables. Thus, the difference in response between the probe and the irrelevants ("probe effect", hereafter) can be used as an index of whether somebody has concealed knowledge about a certain item.

Currently, the CIT is used for forensic purposes only in Japan, but there has been growing interest in it across the world (Matsuda, Nittono, & Allen, 2012). Among the important issues for any potential application of the CIT is the extent to which countermeasures, physical or mental, can reduce its accuracy (Honts et al., 1996). An effective physical countermeasure with polygraphy-based CITs involves pressing one's toes to the floor during the presentation of irrelevants (Honts et al., 1996). Similarly, an effective mental countermeasure entails counting backwards by sevens to irrelevants (Honts et al., 1996). Countermeasures can be effective even with neuroscience-based methods that at first sight may seem more difficult to compromise. One type of countermeasure that can disrupt the accuracy of a common kind of fMRI-based CIT involves associating covert actions with a subset of the irrelevants (Ganis et al., 2011). Its effectiveness is thought to rely on increasing the relative saliency of irrelevants, thus reducing the size of the probe effect.

Although this fMRI study was the first on this topic, it had some limitations. First, it examined only a very specific type of countermeasure, leaving open the issue of whether a more general class of mental countermeasures may be equally or more effective. Second, the countermeasure was applied only to a subset of the irrelevants because applying it to all irrelevants might have artificially increased the saliency of the probe. This may have not only diluted the effect of the countermeasure, as one typically compares the probe with the mean of all the irrelevants, but it may also provide clues of countermeasure use. Thus, countermeasures that can be applied to all items may be more effective. Third, this countermeasure was detectable by examining activation in primary motor cortex, as it involved making irrelevant-specific imperceptible movements with one's fingers and toes and probably engaged motor planning and motor imagery. Finally, no personality dimensions were measured to determine if some people were better than others at implementing the countermeasures.

An interesting class of mental countermeasures that may address the limitations just discussed was tested in recent fMRI work in a different context using standard old/new face recognition paradigms (Rissman, Greely, & Wagner, 2010; Uncapher, Boyd-Meredith, Chow, Rissman, & Wagner, 2015). This work showed that multivariate analyses of brain activation could discriminate well above chance hits (correctly recognized old faces) from correct rejections (correctly rejected new faces) in single individuals. However, the accuracy of the discrimination was reduced to chance by using attentional countermeasures (Uncapher et al., 2015). On the one hand, patterns of brain activation associated with a new face could be made to resemble that of an old (recognized) face by retrieving similar faces already stored in memory and focusing attention to them and by responding as if it was an old face (i.e., pressing the “old” key). On the other, brain activation associated with an old face could be made to look like that of a new face by diverting attention away from the recognition experience and by focusing instead on peripheral perceptual details of the old face and by responding as if it was a new face (i.e., pressing the “new” key).

There are notable differences between standard recognition and CIT paradigms. Specifically, standard recognition paradigms usually employ hundreds of stimuli that are presented only once during study and test and a new stimulus in this paradigm is never encountered before in the study. In contrast, CIT paradigms typically use fewer than 10 stimuli, and these stimuli are repeated tens of times during testing; irrelevant stimuli in these paradigms are not new, in the sense that they have been encountered many times before in the CIT session, like the probes. Thus, at least some of the neural processed involved in discriminating old and new items in standard recognition paradigms are likely to be different from those involved in discriminating between probes and irrelevants in CIT paradigms. Despite these differences, we predicted that attentional countermeasures of this kind would also be effective with CIT paradigms. Thus, we devised a countermeasure that required attention to be focused to superficial probe features and to meaningful memories associated with the irrelevants.

Finally, we tested the hypothesis that more creative individuals may be better at implementing this type of attentional countermeasures. This is because individuals who score high on standard creativity measures have been shown to have more flexible cognitive control (Zabelina and Robinson, 2010), and so they may be better at switching from trial to trial between the two attentional modes required by the countermeasures.

In sum, we compared the probe effect in concealed knowledge, no knowledge, and countermeasure conditions and used both univariate and multivariate analyses to determine the effect of these countermeasures. We expected to find a reliable probe effect in the same prefrontal-parietal network reported in previous CIT studies in the concealed knowledge condition (Ganis et al., 2011; Peth et al., 2015), compared to the no knowledge condition, which in turn was expected to show no probe effect. Furthermore, we expected the probe effect to be smaller in this network in the countermeasure condition. Finally, we expected to find a modulation of countermeasure effectiveness in this network by creativity.

Materials and methods

Subjects

Twenty-three right-handed normal participants (9 females; mean age = 24.2 years) from the University of Padova participated in the study. Exclusion criteria included history or presence of neurological or psychiatric disorders and failure to meet the screening criteria for MRI scanning. Three participants did not follow the instructions and their data were not used. All analyses were carried out on the remaining twenty participants (9 females, mean age = 24.1 years). The Ethics Committee at the University of Padova approved the study and all participants gave signed informed consent. The full subject information and consent form is presented in Appendix 5A.

Stimuli

In each task, the stimuli were six digits (3-8) shown in white against a black background and presented for 750 ms. The stimuli were followed by a black screen with a fixation dot lasting between 1000 and 9000 ms (2000 ms on average), according to a pseudo-random sequence (Dale, 1999). These stimuli were used because they were very similar to each other visually and they had already been successfully used in prior work by this group (Ganis et al., 2016).

Design and procedure

The study was divided into two sessions. The first session took place at the

Neuroradiology Unit, University Hospital of Padua, where fMRI scanning was conducted. After this session, participants were scheduled for a follow up session to complete the Creativity Achievement Questionnaire (CAQ) and the Abbreviated Torrance Test for Adults (ATTA).

Stimuli for the event-related fMRI tasks were presented using E-prime 2.0 software (Psychology Software Tools, Inc., Sharpsburg, PA, USA) and were projected onto the screen of MR-compatible LCD video goggles (VisuaStim XGA, Resonance Technology Inc.) worn by participants. The stimuli were presented at a resolution of 800 x 600 pixels and refreshed at 60 Hz). The three conditions were (i) no knowledge (NK), (ii) concealed knowledge (CK), and (iii) countermeasure (CM). Participants responded on a MR-compatible response box (Evoke Response Pad, Resonance Technology Inc.) using right-handed button presses with their index and middle finger (yes/no, respectively). They were instructed to respond as fast as possible without sacrificing accuracy. There were four runs for each condition and each run consisted of 36 trials where each digit showed up six times. Participants completed a total of 12 runs in the scanner.

All conditions included three types of items (single digits) as follows: (i) one “target” (16.7% of all stimuli) was given participants, who responded truthfully by pressing the “yes” button. This target number was given to participants before starting the fMRI session and it was included to ensure attention was paid in all tasks; (ii) one “probe” (16.7% of all stimuli), whose meaning and response instructions varied by condition. For the NK condition, the probe was just another irrelevant number without any particular meaning for the participants, who simply pressed the “no” key to this item, indicating truthfully they did not know it. No information was given about the probe beforehand and so this control condition simulated the case of participants without concealed knowledge about the probe. In contrast, for the CK condition, participants were given the probe and they were instructed to keep this number secret from the experimenter by pressing the “no” button to pretend they did not know it. Finally, the CM condition was the same as the CK condition, with participants instructed to hide knowledge of the probe by pressing the “no” button. However, in this condition participants were taught to perform a countermeasure on the probe by focusing on superficial aspects of it such as its physical features (e.g., font features, color, size, and so on) as if they were seeing a word but they tried to ignore its meaning by focusing just on how the word looked like on the screen; (iii) four “irrelevants” (66.7% of all stimuli), with no particular meaning for participants, who responded truthfully by pressing the “no” button, indicating they did not know these numbers.

However, in the CM condition, participants were taught to perform a second countermeasure by focusing attention on memories associated with each of the irrelevant numbers (no specific examples were given in order to avoid limiting the kind of associations participants could spontaneously generate). This way, the irrelevants would become more meaningful to participants. The two countermeasures in the CM condition were based on those used by Uncapher and collaborators (Uncapher et al., 2015). Note that in the CM condition too, participants were instructed to respond as fast as possible without sacrificing accuracy. This was done in order to minimize potential differences in the RTs between the CM and the CK conditions.

To ensure that the secret number had no meaning in the NK task and that no systematic countermeasure strategies were applied during the CK task, all participants completed the three tasks in the same order: NK, CK and CM, as in previous work (Ganis et al., 2011). Prior to the fMRI scan, participants underwent a practice session for the NK condition outside the scanner to familiarize them with the stimuli and responses. Instruction and practice for the CK and CM tasks were given before the actual runs while participants were in the scanner. The practice session consisted of 12 trials and was the same for each task but with different instructions.

To enhance the social component of the task, before the CK task participants were told that the main experimenter did not know the secret number, that she was not in the MRI control room, and that she would try to identify the secret number by analyzing their brain images on a computer in a different room. Thus, the job for participants during the CK and CM tasks was to hide their secret number by pressing the "no" button to pretend they did not know it while responding truthfully to all other numbers. In reality, the main experimenter knew the secret number and was in the MRI control room, and this was revealed during the debriefing session at the end of experiment. The debriefing session showed that all participants believed they were hiding their secret number from the experimenter, who was trying to discover it.

Creative Achievement Questionnaire (CAQ, Carson et al., 2005)

We assessed real-world creative behavior with the Creative Achievement Questionnaire, in which participants catalogued any prior creative achievements across ten creative domains (visual art, music, dance, architectural design, creative writing, humor, inventions, scientific discovery, theater and film, and culinary arts). In the music domain, for example, questions range from "I have no training or recognized talent in this area" (score of 0) to "my compositions have been critiqued in a national publication" (score of 7). In the scientific discovery subset, scores vary from "I have no training or recognized ability in this field" (score of 0) to "my work has been cited by other scientists in national publications" (score of 7). Separate domain scores were then combined to form a single index of creative achievement (mean score was 9.45, SD = 8.19, range 2-36). In addition, there is evidence that artists display different personality traits (Feist, 1998, 1999) than creative scientists. Carson et al. (2005) examined the factor structure of the CAQ and provided a two-factor solution identifying an Arts domain (visual art, music, dance, creative writing, humor, and theater and film) and a Science domain (inventions, scientific discovery, and culinary arts). The mean creative achievement score for the Arts domain was 6.00 (SD = 4.41, range 0-14), and for the Science domain was 3.40 (SD = 4.91, range 0-22). These two scores were used to examine whether more creative individuals might be better at implementing countermeasures.

Abbreviated Torrance Test for Adults (ATTA, Goff, 2002):

The ATTA is a shortened version of the Torrance Tests of Creative Thinking (TTCT) with three activities given 3 minutes to respond each. The ATTA provides substantial insight into the creativity of adults by quantifying verbal and figural creative strengths. The creativity index is measured by four norm-referenced abilities (i.e., fluency, originality, elaboration and flexibility) and fifteen criterion-referenced creativity indicators. The mean creativity index was 68.2 (SD = 5.33, range 61-78).

fMRI data acquisition

Whole brain imaging data were acquired using a 1.5T (Siemens Avanto) MRI scanner with an 8-channel head coil. For each participant, functional images were acquired using a gradient-echo planar pulse sequence with 31 axial slices parallel to the anterior-posterior commissural plane, TR = 2000 ms, TE = 30 ms, FOV = 20 cm, FA = 90º, 64 x 64 matrix and 3.125 x 3.125 x 4 mm resolution. During each run 116 functional volumes were acquired, for a total of 464 volumes. A high resolution T1-weighted structural image was acquired, using a magnetization-prepared rapid gradient echo (MPRAGE) sequence with TR = 1900 ms, TE = 2.91 ms, FOV = 25 cm, FA = 8º, 176 sagittal slices, 256 x 256 matrix and 1 x 1 x 1 mm resolution for normalization to a template space. Finally, a T2-weighted structural image co-planar to the functional images was also obtained with 31 axial slices, TR = 7480 ms, TE = 94 ms, FA = 150º, 256 x 256 matrix and 0.8 x 0.8 x 4 mm resolution.

Preprocessing of fMRI data

Brain imaging data were pre-processed and statistically analyzed using SPM8 (Statistical Parametric Mapping, Wellcome Trust Centre for Neuroimaging, London, UK). The first four volumes of each run were not used in the analyses to ensure that T1 equilibrium was reached. For each participant, slice-time and motion correction were applied to the functional volumes, which were then coregistered to the co-planar T2-weighted images and finally with the T1-weighted structural images. T1 images were normalized to the Montreal Neurological Institute (MNI) template using SPM8's segmentation tool. The resulting transformation parameters obtained from the segmentation were applied to the functional images to spatially normalize them to MNI space (2 x 2 x 2 mm voxels for the whole brain univariate analyses, 3 x 3 x 3 mm voxels for the ROI and multivariate analyses). Finally, for the univariate analyses the normalized functional images were spatially smoothed using an 8 mm full-width at half-maximum (FWHM) Gaussian kernel.

Univariate fMRI analyses

For the subject-level analyses, we applied voxel-wise univariate general linear models (GLM) on each participant's preprocessed functional data to obtain individual whole-brain estimates of brain responses to the stimuli presented during the NK, CK and CM conditions. For each condition and run, the GLM included one regressor for each of the 3 types of items (i.e., target, probe, and irrelevant) and six covariate motion parameters. Onset delta functions were convolved with the canonical hemodynamic response function (HRF). Low-frequency noise was eliminated by high-pass filtering at 1/128 Hz.

Contrast images were generated to capture the difference in response between the probe and the irrelevants types for each individual participant (combined across runs) for the NK, CK and CM conditions, which were subsequently used in group-level analyses. The key contrasts were “probe > irrelevants” and “irrelevants > probe”. For the whole-brain group analyses, statistical significance was tested using paired- and one-sample t-tests at $p < 0.001$ at the voxel level (uncorrected), with $p < 0.05$ (FWE-corrected) at the cluster level.

To further quantify the neural response in brain areas known to be involved in this paradigm, we defined spherical ROIs (12 mm radius) around the centers of mass of the seven ROIs reported in Ganis et al. (2011), where the same type of CIT and stimulus timing were used. It was critical to employ ROIs defined in an independent dataset in order to avoid overfitting (Kriegeskorte, Simmons, Bellgowan, & Baker, 2009). The coordinates of the centers of the seven ROIs were: 5, 21, 49 (medial/superior frontal gyrus/anterior cingulate, GFd), 1, -19, 34 (middle cingulate gyrus, GC), 45, 26, -6 (right inferior frontal gyrus/insula, RGFi), -38, 22, -8 (left inferior frontal gyrus/insula, LGFi), 52, -46, 42 (right inferior parietal lobule/supramarginal gyrus, RLPi), -60, -44, 35 (left inferior parietal lobule/supramarginal gyrus, LLPi), and 3, -3, 2 (thalamus, caudate nucleus, lenticular nucleus, Thal). Note that for simplicity in the rest of the paper we will refer to these ROIs by using only the first anatomical structure or its abbreviation. ANOVAs were conducted on the contrast between the probe and the irrelevants with ROI and condition as factors.

Multivariate fMRI analyses

Multivariate analyses were conducted to determine whether patterns of brain activation (multiple ROIs or multiple voxels) could reliably discriminate concealed and no concealed knowledge cases. The multivariate analyses were carried out on spatially normalized contrast images (probe minus irrelevants) without smoothing. Each feature was also normalized across cases by means of a z-score transformation (Hsu, Chang, & Lin, 2003). Classification analyses were carried out with the MATLAB implementation of LIBSVM (Chang & Lin, 2011). Since the number of features far exceeded the number of cases in the multi-voxel analyses, linear support vector machines (SVMs) instead of nonlinear ones were used (Hsu et al., 2003). All analyses reported here were conducted with default cost parameter c = 1. Exploratory analyses with lower and higher values of this parameter (range: 10^{-5} to 10^{5}) showed only small effects on the results and so they will not be reported here. Additional exploratory analyses indicated that quadratic and Radial Basis Function kernels (RBFs) did not lead to better generalization than the linear kernel.

A key issue with multivariate analyses of fMRI data is the high-dimensionality of the datasets, usually requiring data reduction procedures before classification (Jin et al., 2009). To address this issue, in one set of analyses we used as features activation in the 7 ROIs defined in our previous study (Ganis et al., 2011), thus eliminating biases due to selecting features in the same dataset on which the classification is performed. Note that multivariate analyses in this context usually refer to “multi-voxel” analyses, where features are individual voxel activations (Tong & Pratte, 2012). However, they can also encompass “multi-ROI” analyses in which features are ROI average activations. Both multi-ROI and multivoxel analyses (voxels from the ROIs) were performed here.

A recent CIT study (Peth et al., 2015) used signal detection theory methods to quantify accuracy, and showed an Area Under the Curve (AUC) ranging between 0.78 and 0.87 using univariate ROI analyses, and higher classification accuracy (AUC = 0.98) using activation in all grey matter voxels (over 10^5 features) as input to a linear classifier. Although this was unexpected (Jin et al., 2009), we carried out this same analysis on our dataset.

In sum, the multivariate analyses were conducted on three types of data: multi-average ROIs (7 features), multi-voxel ROIs (1069 features), multi-voxel whole brain (26452 features).

For the classification, we used a one-pair-out cross-validation approach in which one pair of cases out of 40 (one case per condition tested, always from different participants) were left out for testing, and training of the classifier was carried out on the remaining cases. Since data in the different conditions were acquired in a within-subject manner, all cases for the left out participants were excluded from the training set as well (e.g., if case CK for participant X was used during testing, then case NK for participant X was excluded also from the training set). This was repeated for all possible 380 pairs of left-out cases (20x19). These analyses were repeated by taking both members of the left-out pair from the same participant, but the differences were negligible, and so they will not be reported.

In applied situations, one might build a classifier on a known set of NK and CK cases and then use it to classify new cases, among which there could be some CM cases.

Thus, in these generalization analyses we trained a classifier to discriminate NK and CK cases, and compared the performance of the classifier on discriminating left-out NK and CK cases and NK and CM cases. The only difference with the previous procedure was that the left-out CK/NK and CM/NK cases were always from 4 different participants, which were not used during training to discriminate CK and NK cases. For these analyses, 10000 random permutations of 4 participants (out of 20) were selected for testing, with training performed on the remaining participants (16 NK and 16 CK cases). The same analyses were repeated by removing the constraint that the left-out cases had to come from different participants, but the results were comparable and so they will not be reported here.

To determine how accurately a classifier discriminated between the different conditions, we used signal detection theory, as detailed in the following section (Peth et al., 2015).

Validity analyses

The validity of behavioural and neural measures in discriminating pairs of conditions was calculated by generating receiver operating characteristic (ROC) curves (National Research Council, 2003) using signal detection theory (Green & Swets, 1966).

This approach provides more complete and precise information than simply calculating accuracy using hits and false alarms at a particular decision value threshold (criterion) that does not reflect directly the distance between the two distributions being discriminated. The key parameter estimated with these analyses is the area under the curve (AUC), which quantifies the separation between two distributions (for example, NK and CK cases) using information from the entire range of criteria (Green & Swets, 1966). An AUC equal to 1 indicates perfect classification accuracy whereas an AUC equal to 0.5 indicates classification at chance. In the present study, we carried out CK vs NK and CM vs NK classifications using the activation probe effect (probe minus irrelevants contrast estimate, for univariate fMRI data) and the decision value distribution for all possible left-out pairs (for multivariate fMRI data). ROC curves for each possible pair of conditions were generated by calculating hits and false positives for criteria spanning the entire distribution of decision values. To determine whether a given AUC value was significantly different from chance (0.5), for univariate fMRI analyses we calculated the 95% confidence interval using parametric methods (Stanislaw & Todorov, 1999). For multivariate analyses, significance was determined using randomization methods (Good, 2005) to empirically estimate the null distribution of AUC values. To calculate the area under the curve (AUC) for the classification between conditions, we used the distribution of decision values resulting from this process (using 'svmpredict' in LIBSVM). The significance of the AUC was determined using a randomization approach as in previous work (Peth et al., 2015). Specifically, we estimated the null distribution of AUCs under the null hypothesis of no difference between conditions, by randomly shuffling the labels of the two conditions and by performing the classification procedure just described. This process was repeated 1000 times. The AUC calculated for the unshuffled data was considered significant at $p < 0.05$ if it was larger than 95% of the values in this null distribution. The difference between the AUC in pairs of conditions was tested for significance by estimating the null distribution of the difference by shuffling the data and by determining whether the unshuffled AUC difference was larger than 95% of the null difference distribution.

Analyses on the effect of creativity measures

To determine if countermeasure use was more effective by more creative individuals, participants were divided into low- and high-creativity subgroups via a median split and probe-irrelevant activation in the right supramarginal gyrus ROI identified in the CM condition was compared between these two groups.

Results

Behaviour

Median target detection accuracy was high in all conditions (NK = 95.8%, CK=91.7%, CM=87.5%), indicating that participants were paying attention to the stimuli. Bootstrap tests comparing the means showed no significant differences in accuracy between any pair of conditions (all *ps* > 0.05).

We conducted two-way ANOVAs comparing RTs between the probe and the irrelevants in the 3 possible pairs of conditions. Follow-up t-tests were carried out to unpack interactions between item and condition.

The ANOVA including the CK and NK conditions (Table 5-1 and Figure 5-1a) showed a main effect of item type, and a trend for a main effect of condition. Importantly, there was a significant interaction between condition and item type. Follow-up t-tests showed that RTs to the probe were slower than to the irrelevants in the CK condition, $t(19) = 4.77$, $p < .0001$, but not in the NK condition, $t(19) = 0.52$, $p = 0.61$, as found in previous studies using similar CIT paradigms (Ganis et al., 2011).

The ANOVA comparing the CM and NK conditions (Table 5-1 and Figure 5-1a) showed main effects of item type and condition, as well as an interaction between condition and item type. Follow-up t-tests revealed that RTs to the probe were slower than to the irrelevants in the CM condition, $t(19) = 3.15$, $p < .01$.

Finally, the ANOVA comparing the CK and CM conditions (Table 5-1 and Figure 5-1a) showed only a main effect of item type, as RTs were slower for the probe than the irrelevants. The mean RTs in the CM condition (554 ms) were numerically slower than those in the CK condition (538 ms), but not significantly so. There was a trend for the RTs to irrelevants (but not probes) to be slower in the CM than in the CK condition, $t(19) = 1.77$, $p < 0.1$ (Figure 5-1).

The same analyses were carried out for error rates (ERs). The ANOVA including the CK and NK conditions (Table 5-1 and Figure 5-1b) showed only a significant interaction between condition and item type. Follow-up t-tests showed a trend for larger ERs for the probe than the irrelevants in the CK condition, $t(19) = 1.85$, $p = .079$, but not in the NK condition. The ANOVA including the CM and NK conditions (Table 5-1 and Figure 5-1b) showed no significant effects. Finally, the ANOVA comparing the CK and CM conditions (Table 5-1 and Figure 5-1b) showed only a main effect of item type, with higher ERs for the probe then the irrelevants.

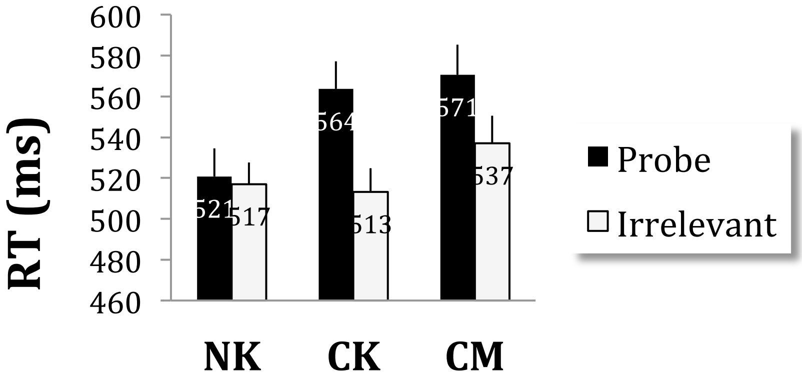
Table 5-1

ANOVAs comparing response times (RTs) and error rates (ERs) for the probe and the irrelevants in all pairs of conditions. For the NK condition, the probe was only nominally defined because participants had no concealed information.

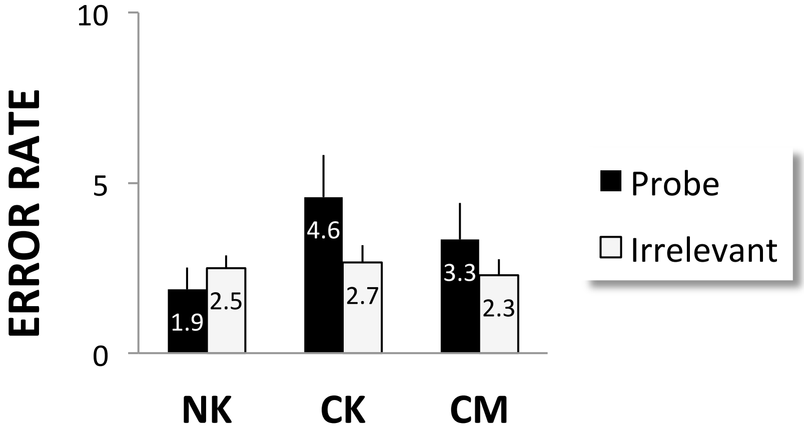
	Conditions								
	CK / NK			CM / NK			CK / CM		
Source (RTs)	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2
Item	21.65	<0.005	0.4	8.09	<0.01	0.30	23.28	<0.001	0.55
Condition	4.16	0.055	0.18	5.59	<0.05	0.23	1.14	0.30	0.06
I x C	22.03	<0.001	0.54	5.46	<0.05	0.22	1.95	0.18	0.09
Source (ERs)									
Item	0.75	0.4	0.04	0.19	0.67	0.01	5.73	<0.05	0.23
Condition	2.64	0.12	0.12	5.59	0.56	0.01	0.71	0.41	0.04
I x C	5.72	<0.05	0.23	2.1	0.16	0.10	0.48	0.50	0.03

Note. Degrees of freedom: 1, 19.

Figure 5-1. (a) Reaction times and (b) error rate for the probe and irrelevants in the NK, CK and CM conditions.



(a)



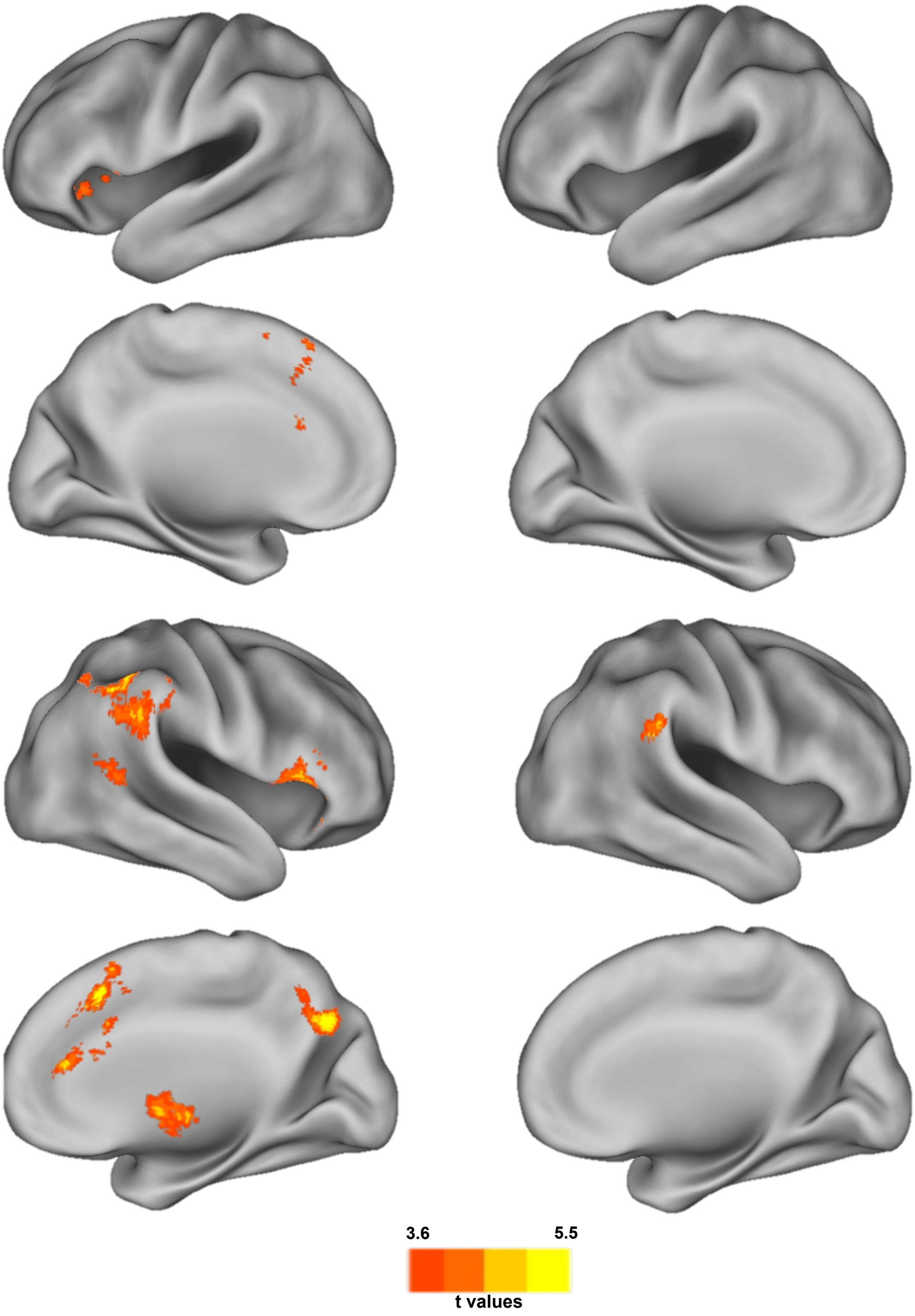
(b)

fMRI: Univariate whole brain analyses

Figure 5-2 shows the results of the whole brain analyses whereas Table 5-2 lists the peak coordinates and minimal t values from the paired t-tests comparing brain activation between the probe and the irrelevants in the NK, CK and CM conditions. In the NK condition, this comparison yielded no significant activation clusters as expected, since participants did not have knowledge of the probe and so it was just another irrelevant. In contrast, in the CK condition, significantly stronger brain activation for the probe then the irrelevants was present in the middle/anterior cingulate cortex and medial frontal gyrus, the bilateral inferior frontal gyri and insula, the right precuneus, the right inferior parietal lobule, and the right caudate nucleus. Finally, in the CM condition the activation was larger for the probe than the irrelevants in the right supramarginal gyrus whereas the opposite pattern was found in the medial orbitofrontal cortex (see Table 5-2, Irrelevants > Probe).

Comparing directly the CK and NK conditions yielded essentially the same regions that were found in the CK condition (Table 5-3), confirming that the effects found in the CK condition were not due to stimulus peculiarities, as these would have been evident also in the NK condition. In contrast, comparing the CM and NK conditions and the CK and CM conditions yield no significant differences.

Figure 5-2. Whole brain map in the contrast of Probe minus Irrelevants in the (a) CK and (b) CM conditions (Top to bottom: left, left-medial, right, right-medial) (Note that for the NK condition, no significant region was found)



(a) CK

(b) CM

Table 5-2

Peak MNI coordinates with Brodmann's Area (BA) and minimal t-statistic of brain regions with significant positive and negative responses in the contrast of Probe minus Irrelevants in the NK, CK and CM conditions. Note that for the NK condition, the probe was only nominally defined because participants had no concealed information.

Condition	Regions	BA	MNI Coordinates			t
			x	y	z	
	Probe > Irrelevants					
NK						
	-	-	-	-	-	-
CK						
	R Middle Cingulate Cortex	24	4	21	37	6.98
	R Anterior Cingulate Cortex	32	6	40	16	5.46
	L Superior Medial Frontal Gyrus	8	-2	26	57	5.24
	L Inferior Frontal Gyrus Pars Orbitalis	48	-46	15	-3	6.63
	L Inferior Frontal Gyrus Pars Triangularis	47	-34	26	-2	4.74
	L Insula	47	-27	21	-3	4.31
	R Precuneus	7	10	-66	36	6.39
	R Inferior Parietal Lobule	40	50	-45	48	5.63
	R Angular Gyrus	39	44	-55	39	5.61
	R Supramarginal Gyrus	48	51	-39	24	5.56
	R Caudate	-	9	3	3	5.43
	R Pallidum	-	15	-3	-2	5.20
	R Inferior Frontal Gyrus Pars Triangularis	45	52	22	9	5.06
	R Inferior Frontal Gyrus Pars Orbitalis	47	48	27	-8	4.86
CM						
	R Supramarginal Gyrus	48	63	-42	27	5.40
	R Supramarginal Gyrus	48	52	-42	25	3.90
	Irrelevants > Probe					
NK						
	-	-	-	-	-	-
CK						
	-	-	-	-	-	-

CM						
	R Gyrus Rectus/Medial Orbitofrontal Cortex	11	2	42	-18	4.75

Note: Significance at all regions for each contrast was tested by a one-sample t-test at $p < 0.001$ at the voxel level (uncorrected), with a significance of $p < 0.05$ (FWE-corrected) at the cluster level.

BA, Brodmann’s area; L, Left Hemisphere; R, Right Hemisphere.

Table 5-3

Peak MNI coordinates and minimal t-statistic of brain regions showing significantly larger probe-irrelevants between conditions (CK/NK, CM/NK, and CK/CM)

Condition	Regions	BA	MNI Coordinates			
			x	y	z	t
CK - NK						
	L Insula	48	-36	17	4	7.04
	L Inferior Frontal Gyrus Pars Orbitalis	48	-45	15	-5	6.63
	L Inferior Frontal Gyrus Pars Oercularis	48	-50	11	1	3.97
	L Superior Medial Frontal Gyrus	8	2	24	45	6.49
	R Inferior Frontal Gyrus Pars Triangularis	45	48	21	9	5.63
	R Superior Frontal Gyrus	10	33	51	10	5.55
	R Pallidum	-	18	-1	0	5.49
	R Caudate	-	15	12	6	4.31
	R Inferior Parietal Lobule	40	48	-45	48	5.44
CM - NK						
	-	-	-	-	-	-
CK - CM						
	-	-	-	-	-	-

Note: Significance was tested with a one-sample t-test at $p < 0.001$ at the voxel level (uncorrected), with $p < 0.05$ (FWE-corrected) at the cluster level.

BA, Brodmann's area; L, Left Hemisphere; R, Right Hemisphere.

fMRI: Univariate ROI analyses

ROI analyses were used to complement the whole brain analyses. We created seven ROIs by defining 12 mm spheres around the center of mass of the ROIs reported in Ganis et al. (2011). These ROIs were used for both the univariate and the multivariate analyses. Figure 5-3 shows the brain activation for the probe effect across all 7 ROIs in the CK, CM and NK conditions.

ANOVAs. We conducted three ANOVAs using the within-subject factors of item type (probe vs. irrelevants), condition (CK vs NK, CM vs NK, and CK vs CM) and ROI (LGFi, LLPi, GC, RGFi, RLPi, Thal, and GFd).

The first ANOVA including the CK and NK conditions showed a main effect of item type and condition (Table 5-4). An interaction between item type and condition indicated that the difference between the probe and irrelevants was larger in the CK than in the NK condition.

The second analysis comparing the CM and NK conditions also showed a main effect of item type and condition (Table 5-4). The interaction between item type and condition indicated that the difference between probes and irrelevants was larger in the CM than in the NK condition.

Finally, the analysis comparing the CK and CM conditions showed a main effect of item type (Table 5-4). Importantly, an interaction between item type and condition indicated that the difference between the probe and the irrelevants was larger in the CK than in the CM condition. Separate follow-up ANOVAs were conducted on the probe and irrelevants separately to unpack the interaction. A main effect of condition revealed that activation to the probe was smaller in the CM than in the CK condition, $F(1,19)=5.19$, $p < 0.05$, $\eta_p^2 = 0.21$. In contrast, although activation to the irrelevants was numerically larger in the CM than in the CK condition, there were no significant effects of condition in the analysis on the irrelevants.

In sum, engaging in countermeasures reduced the size of the probe effect and this was due mostly to a reduced activation to the probe.

Figure 5-3. Brain activation for the probe effect across all 7 ROIs in the CK, CM and NK conditions.

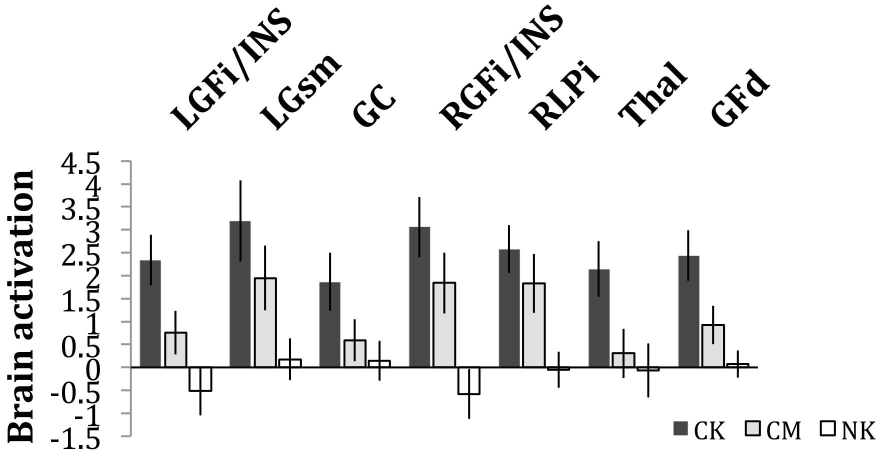


Table 5-4

ANOVAs comparing activation for the probe and the irrelevants across all 7 ROIs and in all pairs of conditions (CK/NK, CM/NK, and CK/CM). Only significant effects involving the factors of interest Condition and Item are shown. Note that for the NK condition, the probe was only nominally defined because participants had no concealed information.

	Conditions								
	CK / NK			CM / NK			CK /CM		
Source	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2
Item	14.28	<0.001	0.43	4.89	<0.05	0.21	26.26	<0.001	0.58
Condition	14.88	<0.001	0.44	7.71	<0.05	0.29	2.99	0.1	0.14
I x C	21.99	<0.001	0.54	6.36	<0.05	0.25	6.08	<0.05	0.24

Note. Degrees of freedom: Item: 1,19; Condition: 1,19; IxC: 1,19

Countermeasures and creativity

Individuals who scored high on the CAQ art domain showed a smaller probe effect in the right supramarginal gyrus (rSMG) ROI identified in the CM condition than individuals who scored low 1.24 vs 4.35, respectively, $t(14) = 2.64$, $p = .019$. In other words, more creative individuals had a smaller probe effect in the CM condition. No effect was found for the science domain, 2.05 vs 1.73, $t(13) = 0.25$, $p = .807$, and the ATTA, 2.07 vs 3.39, $t(14) = 1.07$, $p = .302$.

Potential evidence of countermeasure use

In the CM condition, we instructed participants to associate the irrelevants with stored autobiographical memories. Thus, we predicted that brain regions involved in memory retrieval would exhibit more activation when processing irrelevants than the probe in the CM condition (reverse effect). To this aim, we carried out an exploratory analysis (i.e., $p < .001$ at the voxel level with an extent threshold of 10 voxels) to investigate whether irrelevants activated memory-related brain regions more than probes. This contrast revealed greater brain activations in a left parahippocampal gyrus (PHG) cluster (-28, -30 -17) with an extent of 45 voxels. A nearly identical region, left Hippocampal gyrus/PHG (-24, -33, -17), was also identified in a previous meta-analysis on autobiographical event memory studies (McDermott, Szpunar, & Christ, 2009). In order to examine whether this reverse effect only existed in the CM condition, we also extracted the response estimate from the standard CK condition to compare the difference. The result showed that the reversed effect for the PHG was higher in the CM (*M* = 1.06, SD = 0.46) than the CK (*M* = -0.34, SD = 0.40) condition, *t*(19) = 2.493, *p* < .05, *d* = 0.56.

Validity

Univariate ROI analyses. It was possible to discriminate above chance between CK and NK cases using the average activation from any of the seven ROIs (Table 5-5). The AUC varied among ROIs, and it was highest for the left inferior frontal gyrus (0.85), and lowest for the thalamus (0.73). In contrast, it was possible to discriminate CM and NK cases above chance only using activation in the right inferior frontal gyrus, the right inferior parietal lobule, and the left inferior frontal gyrus (0.74, 0.71, and 0.69, respectively). The countermeasures resulted in significantly smaller AUC values for the left inferior frontal gyrus, the middle cingulate gyrus, the thalamus, and the medial frontal gyrus.

The same analyses carried out on the average of the seven ROIs (after score normalization) showed an AUC of 0.86 for discriminating CK versus NK cases, significantly larger than the AUC of 0.71 found for discriminating CM versus NK cases (J. A. Hanley & McNeil, 1983).

Table 5-5

Area under the curve (AUC) for the comparisons CK vs NK and CM vs NK for 7 ROIs

	ROI							
Comparison	LGFi	LLPi	GC	RGFi	RLPi	Thal	GFd	Mean
CK vs NK	**0.85***	**0.76**	**0.74***	**0.82**	**0.81**	**0.73***	**0.83****	**0.86***
CM vs NK	**0.69**	0.66	0.57	**0.74**	**0.71**	0.56	0.61	0.71

Note: LGFi: left inferior frontal gyrus/insula, LLPi: left inferior parietal lobule/supramarginal gyrus, GC: middle cingulate gyrus, RGFi: right inferior frontal gyrus/insula, RLPi: right inferior parietal lobule/supramarginal gyrus, Thal: thalamus/caudate nucleus/lenticular nucleus, GFd: medial frontal gyrus/superior frontal gyrus. AUC numbers in bold indicate values above chance (0.5). The asterisks indicate a significant difference between the CK/NK and CM/NK AUCs: * p<.05, ** p<.005.

Multiaverage ROI analyses. In this analysis, average activation in each of the seven ROIs used in the univariate analyses was employed as a feature for a linear SVM classifier. The AUC for discriminating CK and NK cases was 0.85. The AUC for discriminating CM and NK cases was significantly smaller, 0.63, and not statistically different from chance.

Multivoxel ROI analyses. In this analysis, activation in a total of 1069 voxels from the 7 ROIs (unsmoothed data) was used as input to a linear SVM classifier. The AUC for discriminating CK and NK cases was 0.83, whereas that for discriminating CM and NK cases was significantly smaller, 0.63, and not statistically different from chance.

Multivoxel whole brain analyses. In this analysis, activation in a total of 26452 gray matter voxels (unsmoothed data) was used as input to a linear SVM classifier. The AUC for discriminating CK and NK cases was 0.80, whereas that for discriminating CM and NK cases was 0.79.

Generalization analyses

In these analyses a classifier was trained to discriminate NK and CK cases and then tested on discriminating left-out CK/NK and CM/NK cases.

Multivariate ROI analyses. The AUC for discriminating CK and NK cases was significantly larger than for discriminating CM and NK cases, 0.84 and 0.73, respectively. Both AUCs were significantly different from chance. At a fixed false alarm rate of 20%, the countermeasures reduced hit rates from 80% to 61%.

Multivoxel ROI analyses. The AUC for discriminating CK and NK cases was significantly larger than for discriminating CM and NK cases, 0.84 and 0.68, respectively. Both AUCs were significantly different from chance. At a fixed false alarm rate of 20%, the countermeasures reduced hit rates from 75% to 37%.

Multivoxel whole-brain analyses. The AUCs for discriminating CK versus NK cases and CM versus NK cases were 0.79 and 0.72, respectively. Both AUCs were significantly different from chance, but they were not different from each other. At a fixed false alarm rate of 20%, the countermeasures reduced hit rates from 65% to 42%.

Discussion

This study found that activation in a set of prefrontal, parietal and subcortical regions differentiated between the probe and irrelevants in a concealed knowledge condition, confirming and extending the results of previous CIT work (Ganis et al., 2011; Peth et al., 2015). Such differences were still present when comparing them to a matched no-knowledge condition, ensuring that they were not due to stimulus peculiarities . Critically, the mental countermeasures tested in this study reduced the size of the probe effect and decreased classification accuracy, extending the countermeasure findings of previous fMRI work (Ganis et al., 2011).

Brain regions involved in the probe effect

The pattern of brain activation for the probe effect was comparable to that found in previous fMRI CIT studies (Ganis et al., 2011; Peth et al., 2015), as discussed next.

First, the probe engaged the VLPFC (bilaterally) and the adjacent anterior insula (in the left hemisphere) more than the irrelevants. The substantial heterogeneity in the functional organization of these frontal regions, along both the rostro-caudal and the laterality dimensions, has made it difficult to determine their precise role in cognition (Levy & Wagner, 2011). One proposal is that the VLPFC and adjacent insula are involved in reflexive orienting of attention to behaviorally relevant changes in the environment. Indeed, these regions, together with parts of the medial prefrontal cortex, have been considered a key component of a salience network (Seeley et al., 2007) and they have also been conceptualized as the frontal nodes in a largely right-lateralized ventral attentional network that includes the right inferior parietal cortex as well (Corbetta & Shulman, 2002). Another proposal has focused instead on the role of these frontal regions in motor inhibition processes (Aron et al., 2004; Swick, Ashley, & Turken, 2008), as they are consistently engaged for instance by Go/No-Go tasks. Meta-analytic approaches have shown that different subregions of the ventrolateral prefrontal and adjacent insular cortex tend to respond differently to tasks that tap into attentional reorienting and motor inhibition processes, though the segregation is not clear-cut (Chang, Yarkoni, Khaw, & Sanfey, 2013; Levy & Wagner, 2011). For instance, the pars opercularis of the right inferior frontal gyrus tends to be engaged by motor inhibition but not attentional reorienting tasks whereas the inferior frontal junction and the anterior insula tend to be engaged bilaterally (but with a right hemisphere bias) by both attentional reorienting and motor inhibition tasks (Change et al., 2013; Levy & Wagner, 2011). In addition to attention reorienting and motor inhibition tasks, the VLPFC is usually engaged by several other classes of tasks as well. Especially relevant for the CIT are the potential roles of the VLPFC in memory processes such as encoding and retrieval (e.g., Fletcher et al., 2002; Iidaka, Sadato, Yamada, & Yonekura, 2000) and in social cognitive processes such as action imitation (e.g., Levy & Wagner, 2011; Molnar-Szakacs, Iacoboni, Koski, & Mazziotta, 2005).

One or more of these processes could account for the concealed knowledge probe effect in VLPFC and the insula in CITs. Indeed, the probe is more salient than the irrelevants (since the probe is the only item associated with the crime episode and it is presented infrequently) and lying to the probe is likely to require inhibiting a prepotent truthful response (Verschueure, Ben-Shakhar, & Meijer, 2011). Furthermore, the probe in the CK condition should engage retrieval processes more strongly than irrelevants because it is associated with a pre-experimental episode. Finally, it is also likely that the probe engages social cognitive processes more strongly than the irrelevants because deception only occurs for the probe and often instructions mention that a judge would try to detect deception based on various deception cues.

The CIT literature has been mixed about the interpretation of the role of the VLPFC. This is in large part because CIT paradigms usually engage more than one of the processes just discussed and designing paradigms that isolate individual processes has proven to be very difficult. For instance, an fMRI study that attempted to eliminate response competition processes in the CIT interpreted VLPFC activation as reflecting memory-related processes (Gamer, Klimecki, Bauermann, Stoeter, & Vossel, 2012). Results from more recent work, however, have been interpreted as indicating that response competition processes are critical for VLPFC activation (Suchotzki, Verschuere, et al., 2015). A detailed discussion of why such discrepancies might exist goes beyond the scope of this paper, but the main point is that they are probably due to the difficulty in isolating individual processes in CIT paradigms.

Second, the probe engaged medial prefrontal cortical regions more than the irrelevants, including parts of the middle and anterior cingulate and the superior portions of the medial frontal gyrus. Portions of these medial prefrontal regions are activated by the same attentional and response inhibition tasks that engage the VLPFC and have been implicated in monitoring the conflict between competing responses (Braver, Barch, Gray, Molfese, & Snyder, 2001; Rushworth, Walton, Kennerley, & Bannerman, 2004).

Third, portions of the right inferior parietal lobule, angular gyrus, and supramarginal gyrus were also recruited more strongly by the probe than the irrelevants. These regions overlap in large part with the right temporo-parietal junction (TPJ), which has been implicated in attentional reorienting in a number of domains (Corbetta & Shulman, 2002; Shomstein, 2012). More recently, the TPJ (bilaterally) has been suggested to be involved in contextual updating, that is in updating an internal model of the physical and social environment in order to revise expectations and responses after detecting a change in the environment (Geng & Vossel, 2013). Note that in previous CIT work the laterality of activation in the inferior parietal lobule has varied, with some studies reporting only activation in the right hemisphere (e.g., Gamer et al., 2012; Peth et al., 2015) and others in both hemispheres (e.g., Cui et al., 2014; Ganis et al., 2011), probably due to differences in stimuli and paradigms.

Fourth, there was also a region that has not been reported in previous CIT studies, the right precuneus. Using these coordinates in the Neurosynth database (Yarkoni, Poldrack, Nichols, Van Essen, & Wager, 2011) suggests a link with successful retrieval, which would make sense for the probe. Also, it makes sense that it is not as active in CM because the probe countermeasure attempts to make retrieval less likely. However, since this region was not found in previous CIT studies, we could not use it as an apriori region for the classification analyses.

Classification accuracy (CK vs NK)

With the exception of the precuneus, the pattern of activation corresponds closely to that found in our previous study (Ganis et al., 2011). Thus, the 7 ROIs identified in that study were used to test the accuracy of single participant classification in the current study. Results showed that the average activation in each of the 7 ROIs could be used to classify CK and NK cases (activation difference between the probe and the irrelevants) well above chance in the current study. As found in the previous study, the ROIs with the highest accuracy were the left and right ventrolateral prefrontal cortex (AUC = 0.85 and 0.82, respectively), and the medial frontal cortex (AUC = 0.83). However, the accuracy rates were lower than in our previous study in which an AUC of 1 was found (Ganis et al., 2011). This may be due to a number of factors. First, a 1.5T scanner at a different site was used in this study, which most likely produced data with a lower signal to noise ratio. Second, the stimuli used in this study (single digits) were much less salient than those used in the previous study (one's date of birth). Third, the analyses were different, as this study used beta values rather than number of significant voxels within a region, as input to the classifiers. Fourth, the population used in this study (University of Padova students and affiliates) was more varied than that used in the previous study (Harvard University undergraduates), which may have increased the variance in the data.

The univariate ROI accuracy rates found in this study are similar to those found in the study by Peth and collaborators (Peth et al., 2015), where they varied between .66 and .87, depending on the compared conditions. Similar accuracy rates were also found with the multivariate ROI analyses. Note, however, that the results of these two studies are not directly comparable because our study did not classify the probe and irrelevants in various conditions (Peth et al., 2015), but rather it classified the probe effect (the difference between the probe and the irrelevants for each participant) between conditions (NK, CK, and CM). Classifying the probe effect across conditions should help reduce potential individual differences in brain responses by subtracting activation to irrelevants from that to the probe.

A recent study used multivariate analyses with linear SVMs on data from a CIT paradigm and reported an AUC of .98 for classifying probes in a concealed knowledge group and irrelevants in a no-knowledge group by employing 125,570 grey matter voxels (whole brain data) without any data reduction (Peth et al., 2015). Interestingly, classification accuracy using whole brain data was not consistently high because the AUC for classifying the probe and irrelevants in a guilty intention group was only 0.71, numerically lower than with univariate ROI analyses, and not significantly different from chance. This suggests that idiosyncracies in the datasets may play an important role in the results. Furthermore, between-cell statistics were not reported to show that whole brain analyses were significantly more accurate than the corresponding univariate or multivariate ROI analyses. For example, the AUC for univariate analyses using combined ROIs was .87 (Guilty Action probe versus Guilty Intention irrelevants), which is probably not significantly higher than the AUC of .90 found for the corresponding whole-brain multi-voxel analysis (Peth et al., 2015).

In the current study, whole brain multivoxel analyses did not improve classification accuracy of CK versus NK cases compared to multivoxel ROI analyses, but rather they slightly decreased it, from .83 to .80, a statistically non-significant difference. A potential explanation for this is that the 7 ROIs already included voxels that maximally discriminated between the CK and NK conditions (given that they were defined in a previous study contrasting these two conditions) and so including other voxels added more noise than additional information useful for the classification, resulting in a slightly degraded performance.

Effect of countermeasures

The VLPFC, insula, and medial prefrontal regions discussed earlier have very high base rates, which means that they are engaged by many different tasks (Yarkoni, Poldrack, Nichols, Van Essen, & Wager, 2011). An important consequence of this is that CIT methods that rely on activation in these regions are likely to be vulnerable to mental countermeasures because many cognitive processes unrelated to concealed information can be used to modulate neural responses in these regions. Indeed, the physical countermeasures used in the previous study (Ganis et al., 2011) substantially reduced classification accuracy using mean activation within ROIs. In this study, mental/attentional countermeasures reduced the size of the probe effect across the 7ROIs reported in Ganis et al., (2011), relative to the CK condition, mostly by decreasing activation elicited by the probe. This shows that the probe countermeasure was successful at reducing the saliency of the probe. Activation to irrelevants became numerically larger as a result of the corresponding countermeasure, but not significantly so. This may be due in part to the nature of saliency itself, which is inversely related to frequency of occurrence. It is possible that the countermeasure on the one hand made each irrelevant more salient by virtue of the associated memory, but on the other hand, in trying to make all irrelevants more salient, it results in a net decrease in saliency for all irrelevants. It is worth noting that although the stimuli (Ganis et al., 2011 used dates) and population used were different, the brain regions involved in the probe effect corresponded with each other. This shows that it is promising to use CIT paradigm as a lie-detection tool as similar neural activities are generated regardless of the stimuli and population.

The countermeasures strongly affected classification rates, relative to CK/NK classification, in both univariate and multivariate ROI analyses. For the univariate analyses, accuracy decreased from .86 to .71 when combining ROIs. For the multiaverage ROI analyses, accuracy decreased from .85 to .63, whereas for the multivoxel ROI analyses it decreased from .83 to .63. However, relative to CK/NK classification, the countermeasures had virtually no effect on the AUC based on whole-brain multivoxel analyses, 0.8 and 0.79, respectively. In contrast, relative to multivoxel ROI analyses, multivoxel whole-brain analyses improved classification accuracy for countermeasures, from .63 to .79. This result, and the opposite pattern described earlier for CK/NK classification, may be explained by two observations: i) the 7 ROIs already included voxels that maximally discriminated between the CK and NK conditions (given that they were defined in a previous study contrasting these two conditions) and so including other voxels added more noise than additional information useful for the classification, resulting in a slightly degraded performance, and ii) voxels that maximally discriminated between CM and NK cases were not included in the 7 ROIs, and so adding other regions improved classification accuracy for CM/NK classification.

To test this idea, we ran a batch of multivoxel analyses that included the 7 original ROIs and gradually added more and more spherical ROIs (up to 200, radius = 12mm) centered at random locations. Results showed that, as the number of additional ROIs grew to 200, CK/NK classification accuracy slightly declined (gradually going from .83 to .80), whereas CM/NK classification accuracy gradually increased from .63 to .79. This suggests that most information useful for the CK/NK classification is already available within the original ROIs, whereas additional information about the CM/NK classification is distributed across voxels outside the original ROIs.

This interpretation was also confirmed by a multivoxel analysis carried out on all voxels not included in the 7 ROIs (radius = 20 mm). Results showed that, compared to using only the 7 ROIs, CK/NK classification accuracy decreased from .83 to .74 whereas CM/NK classification accuracy increased from .63 to .70. In other words, collectively, there was more information useful for the CM/NK classification in voxels outside than inside the 7 ROIs, whereas the reverse was true for the CK/NK classification.

With regard to the reverse effect (irrelevants > probe), the activation in the parahippocampal gyrus (PHG) is consistent with the idea that this region was engaged in episodic retrieval and that participants were performing the countermeasure as instructed. Uncapher et al. (2015) used a similar cognitive countermeasure in an fMRI study where participants were asked to focus on perceptual aspects of previously seen faces, "old" faces, (the equivalent of the countermeasure applied to the probe item in the current study) and to prevent memories for such faces from coming to mind; conversely, for new faces participants were instructed to bring to mind a known individual that resembled those faces and relive any memories associated with the individual (the equivalent of the countermeasure applied to the irrelevant items). The reverse effect was also found in their study where bilateral hippocampus and left angular gyrus (AnG), typically engaged during memory retrieval, showed greater activity for new than old faces. Both studies show that cognitive strategies can greatly affect the ability of fMRI-based classifiers to detect memories. It is also worth noting that Uncapher et al. (2015) instructed participants to apply countermeasures for entire 10-second trial whereas, in the current study participants were only instructed to carry out countermeasures for the duration of each stimulus (750 ms). Even this brief countermeasure, however, was enough to interfere with fMRI-based detection of concealed knowledge.

Countermeasures and creativity

Neural activity in the right supramarginal gyrus (rSMG) identified in the CM condition was found to be associated with creativity. The result demonstrated that higher creative individuals tended to implement countermeasures more effectively by showing a smaller probe effect. As hypothesized, more creative individuals were better at implementing this type of attentional countermeasure due to more flexible cognitive control as found in Zabelina and Robinson (2010). Flexible cognitive control is defined in terms of a smaller target congruency effect after incongruent primes and a larger target congruency effect after congruent primes (Gratton, Coles, & Donchin, 1992). Zabelina and Robinson (2010) assessed individual differences in flexible cognitive control in a basic color-word Stroop task and found that more creative individuals displayed greater modulation of the cognitive control system across trials. The instructions of the CM condition required participants to apply the countermeasures to probe and irrelevants differently; thus, participants had to switch rapidly between these two processing modes. In line with the previous study, the current result demonstrated that higher creative individuals were more capable of switching attention rapidly by showing a smaller probe effect in the CM condition.

With regard to the rSMG, it has been found to be involved in attentional reorienting along with neighboring regions, i.e., inferior parietal lobule, angular gyrus and TPJ (see the earlier discussion). Furthermore, previous fMRI studies have shown a negative link between parietal brain activity and creativity. For example, Berkowitz and Ansari (2010) found that musicians deactivated the right TPJ during melodic improvisation while non-musicians showed no change of activity in this region. They suggested that the deactivation of the rTPJ might allow the experienced musicians to enter a more focused attentional state. Also, the deactivation of the right TPJ, including SMG, has been postulated to occur in response to top-down modulation during goal-driven behavior in order to inhibit attentional deviation (Corbetta, Patel, & Shulman, 2008; Shulman et al., 2003; Todd, Fougnie, & Marois, 2005). Taken together, we proposed that smaller probe effect in the rSMG when applying countermeasures was associated with attentional switching and its activity was modulated by creative thinking.

Potential limitations

This study has a number of limitations. First, the paradigm and stimuli were intentionally minimalistic in order to reduce potential perceptual difference between stimuli that can be problematic for the interpretation of CIT results (Ganis et al., 2016). This approach was justified here because: i) the main goal of the study was to determine whether the accuracy of 3S CIT paradigms is affected by attentional countermeasures, regardless of actual accuracy rates in the field, and ii) even elaborate mock crime scenarios are generally far from ecologically valid situations because participants still know that the entire situation is fictitious, they are just following instructions, and the items employed in the scenarios don't usually have any intrinsic value to the participants.

Second, the tasks were administered in a within-participant manner, as in our previous study (Ganis et al., 2011). To eliminate the effect of potential within-subject correlations during testing, in the classification analyses we always used test cases from participants that had been excluded from the training set. For example, if the case NK for a participant was used for testing, also the CK and CM cases for that participant were excluded from training. The analyses were repeated by removing this constraint as well, but the results were comparable.

Third, although the number of cases (20 per condition) was comparable to that used in previous studies, this is still a relatively small number to assess classification accuracy, especially given the large number of features used in the multivariate classification analyses.

Conclusions

Together with Study 3, memory and attentional countermeasures were shown to be effective in reducing the probe effect in neuroimaging-based CIT paradigms. These two studies provide evidence that further research is needed before neuroimaging methods are sufficiently robust to be used to detect concealed information accurately in the real situation. Given these results, the lack of a difference between the activation of probes and irrelevants in a typical CIT paradigm does not imply that participants have no knowledge about the probe. Indeed, the result could be a false negative because of countermeasures applied by suspects who have actually committed the crime. These two studies also support the general point that the vulnerability of the neuroimaging methods for lie detection is due to poor specificity. The lack of distinct neural activity patterns that are selective for deception makes it relatively easy to alter such patterns using countermeasures.

CHAPTER 6

GENERAL DISCUSSION

Summary of Findings

These four studies advance our understanding of deception by examining the underlying cognitive processes and neural mechanisms and the validity of the lie detection paradigms. The first behavioural study determined whether creative cognition contributes to the individual differences in deceptive abilities within a socially interactive setting. Results of a multiple regression analysis showed that the ability to generate lies was predicted by higher creative cognition and by the number of uncued lies told when given a choice. This result provides the first evidence for a link between deceptive abilities and creative cognition. This study also provided evidence that lie generation and detection are independent abilities by showing no correlations between them and a different pattern of correlation with creative cognition. A second ERP study investigated the neural mechanisms underlying the generation of uncued lies using a novel bluffing paradigm. Results showed response-locked ERP differences just before the response between uncued lies and truths, but no stimulus-locked differences. This study suggests the engagement of additional cognitive processes during lying, related to holding both lies and truths in working memory and to moral conflict. Finally, parallel fMRI and ERP studies examined the validity of the CIT paradigm when mental countermeasures were applied and determined the role of creativity in using this type of countermeasures. Results showed that this type of countermeasures requiring attentional switch degraded the neural signatures of deception (i.e., hiding information) and the fMRI study found that this effect was modulated by creative cognition. These two parallel studies suggest that more research is needed for developing the lie detection methods before they can be applied to real life situations.

The Difference in P300 between the Bluffing and CIT Paradigms

The amplitude of stimulus-locked P300 was measured in both bluffing (Study 2) and CIT (Study 3) paradigms. However, the amplitude of the P300 associated with lying showed an opposite pattern in these two paradigms. In the bluffing paradigm, although there was no P300 difference between uncued lie and uncued truth, a decreased P300 for lying was found compared with cued truth. This is consistent with previous studies using cued conditions (Hu et al., 2011; Johnson et al., 2003, 2005; Pfister et al., 2014; Suchotzki et al., 2015; Wu et al., 2009), and the result was interpreted as an indication of the dual task character of lying: lying requires higher cognitive load in order to keep the truth active, monitored and inhibited at the same time. In contrast, the ERP study using the CIT paradigm observed a larger P300 for lying (i.e., the probe) compared with truth-telling (i.e., the irrelevants). This P300 is elicited by rare stimuli (16.7%) embedded within a series of standard stimuli (66.7%), and its amplitude is inversely related to the subjective frequency of the eliciting stimulus (Donchin & Coles, 1988), and increases with the level of significance of the stimulus (Berlad & Pratt, 1995; Johnston et al., 1986). The unequal proportion of lying and truth-telling trials makes the oddball effect override the cognitive load effect of deception in the CIT paradigm. This difference of P300 results between these two deception paradigms also demonstrated that no specific neural activity pattern is associated with lying; instead, neural activities vary depending on paradigm design. This supports the general point that the neuroimaging methods for lie detection lack specificity and paradigm designs should carefully be considered when interpretating the results.

Countermeasure Effects with ERP and fMRI Methods

The same mental countermeasures were applied in the ERP (Study 3) and fMRI (Study 4) studies using the same CIT paradigm. Both results showed that the countermeasures successfully reduced the probe effect and the accuracy to discriminate between guilty and innocent cases was lower. However, when we examined the effect on the probe and irrelevants separately, the results were not fully consistent with each other. The fMRI study showed that the activation to the probe was significantly smaller in the condition where countermeasures (i.e., the CM condition) were applied than in the standard CIT condition (i.e., the CK condition), whereas activation to the irrelevants was not significantly larger in the CM than in the CK condition. Thus, the smaller size of the probe effect was due mostly to reduced activation to the probe. On the other hand, although the same activation pattern was found in the ERP study (i.e., smaller activation to the probe and larger activation to the irrelevants in the CM than in the CK condition), only the difference for the irrelevants approached significance ($p = .072$). In other words, reducing the size of the probe effect was due more to an enhanced activation to the irrelevants in the ERP study. This inconsistency may be due to some reasons.

First, the data acquisition time frame is different between ERP and fMRI methods. ERPs measure the electrical activity of the brain and can detect changes with millisecond temporal resolution: ERP results reflect brain activity occurring within a second after a stimulus or a response. On the other hand, fMRI measures brain activity based on much slower BOLD activity, which reflect changes in regional cerebral blood flow. The hemodynamic response lasts over 10 seconds and peaks at 4 to 6 seconds.

FMRI results capture brain activity over a few seconds after a stimulus or a response. It is possible that when seeing the probe, it took some time for top-down modulation (i.e., focus on the physical features of the stimulus) to inhibit bottom-up processes (i.e., stimulus-driven response). Thus, fMRI results showed the activation to the probe was significantly smaller in the CM than in the CK condition, but not ERP results.

Second, the size of the P300 probe effect in the ERP study relies on the saliency of the probe. Participants were instructed to make the irrelevants more salient by associating the items with meaningful memories. Due to the characteristic of the P300, it may be easier to capture the initial saliency upon seeing the irrelevants. Thus, stronger activation to the irrelevants in the CM than in the CK condition approached significance. In contrast, fMRI results captured brain activation for a few seconds, and it required participants' constant effort to make the memory associations. Participants might not be able to focus attention on doing this countermeasure consistently for each irrelevant item throughout the task, resulting in no significant difference of the activation to the irrelevants between the CM and CK conditions in the fMRI study.

Third, the factors just mentioned may be further amplified by the timing differences in stimulus presentation in the ERP and fMRI studies. Indeed, stimuli were presented much more slowly and with much higher temporal variability during the fMRI than the ERP study in order to enable proper deconvolution analyses. It is possible that such differences in presentation rate regimes may have affected the implementation of the countermeasures used for the probe and the irrelevants.

Given that countermeasures affect the validity of each neuroimaging method in a slightly different way, it is important to examine different types of countermeasures with various methods in order to understand the whole picture of how such countermeasures affect the validity of these methods. With such knowledge, we will be more likely to detect if participants apply any countermeasures in a real scenario. On the other hand, Johnson (2014) suggests that a promising direction in the future may be to incorporate multiple, simultaneously and sequentially obtained behavioural and different brain-based measures to capture the whole picture of deception-related processes from different angles. With multiple measures, the classification between guilty and innocent individuals would be more accurate.

Limitations and future work

The work conducted in this thesis has some limitations. First, an ideal plan was to conduct the ERP study (Study 2) that could examine the neural processes underlying real-life deception in a similar interactive deception setting to the first behavioural study

(Study 1). For example, provide a circumstance for participants to justify their claims or to describe their previous events and feelings instead of only allowing them to speak one word as a response. By doing so, we would be better at characterizing the behavior-brain relationships by bridging these two studies. Also, we might be able to replicate the behavioral results, e.g., creative cognition can facilitate the ability to generate lies. However, due to the constraints of using EEG caps and electrodes, we had to compromise on the paradigm and adapt it to an ERP-friendly design. Future research may use fMRI methods to investigate the lie generation process similarly to the setting used in Study 1. That would provide more information about how brain regions are involved in the lie generation process and how creative cognition modulates the brain activity to enhance this ability.

Second, the same countermeasures were applied in a typical concealed information paradigm but using ERP and fMRI methods in Study 3 and Study 4, respectively. Although both studies revealed the same conclusion that these mental countermeasures were efficient in decreasing the validity of the test, using different groups of participants limited our ability to compare the two parallel studies. A within-subject design might enable to address the following questions. First, can a participant who successfully misleads one method (e.g., ERP) also mislead another (e.g., fMRI)? This can also answer whether the outcomes from these two neuroimaging-based methods are consistent or not. Second, are there any individual differences in the ability to apply the countermeasure to the probe or to the irrelevants, respectively? It is possible that some people are better at applying the “probe countermeasure” and some people are better at applying the “irrelevants countermeasures”. We could test this by observing whether participants are consistently doing better at one particular countermeasure in these two studies and whether it is associated with any cognitive abilities, e.g., working memory span, inhibition, attention, etc. Third, do multivariate analysis combing ERP and fMRI results achieve better accuracy rate to discriminate guilty and innocent cases? If the same group of participants took part in both ERP and fMRI studies, we would be able to incorporate the data from these two methods and conduct validity analyses. By doing so, we could examine whether this combined classifier is better than any single neuroimaging method. However, due to the lack of access to the MRI facility in the University of Plymouth, this ideal plan could not be carried out. Future research can address this idea.

Conclusions

This thesis advances understanding of deception from both theoretical and applied angles by investigating the cognitive processes and neural mechanisms underlying deception using various behavioural and neuroimaging methodologies and examining the cognitive strategies that can compromise the validity of neuroimaging-based classifiers to detect deception. Overall, this series of studies demonstrates a new approach to investigate the individual differences in deceptive communication by showing the association with creativity, and suggests that the paradigm design is crucial to study real-life deception in a laboratory setting. Furthermore, it shows that neuroimaging methods to detect deception are vulnerable to mental countermeasures due to poor specificity. Thus various countermeasures need to be assessed before the neuroimaging methods can be used in applied settings.

APPENDIX A

Ethical approval of research

PLYMOUTH UNIVERSITY
FACULTY OF HEALTH & HUMAN SCIENCES

Application No:

(for FREC use)

RESEARCH
WITH
PLYMOUTH
UNIVERSITY

Faculty Research Ethics Committee
APPLICATION FOR ETHICAL APPROVAL OF RESEARCH

Title of research:

Creativity in Deceptive Communication

1. Nature of approval sought (Please tick relevant boxes)

(a) PROJECT*: ☐ (b) PROGRAMME*: ☒

If (a) then please indicate which category:
- Funded research project ☐
- MPhil/PhD project ☒
- Other (please specify): ☐

**Note: In most cases, approval should be sought individually for each project. Programme approval is granted for research which comprises an ongoing set of studies or investigations utilising the same methods and methodology and where the precise number and timing of such studies cannot be specified in advance. Such approval is normally appropriate only for ongoing, and typically unfunded, scholarly research activity.*

2. Investigators/Supervisors

Principal Investigator (staff or postgraduate student)*:
Name: Chun-Wei Hsu
Email: chun-wei.hsu@plymouth.ac.uk

Address for written correspondence:

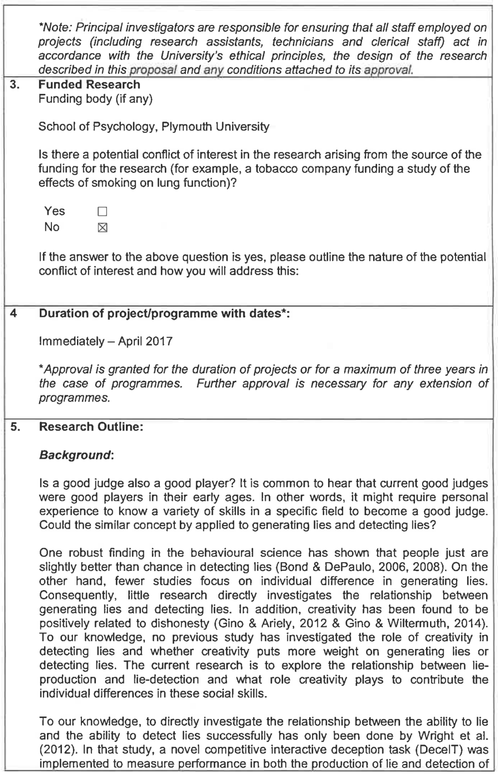
School of Psychology, Link Building, Room 301
Plymouth University
PL4 8AA Plymouth

Other staff investigators:

Director of Studies/other supervisors (only where Principal Investigator is a postgraduate student):

Dr. Giorgio Ganis, Dr. Haline Schendan, Dr. Michaela Gummerum.

Please indicate Department of each named individual, including collaborators external to the Faculty: School of Psychology



**Note: Principal investigators are responsible for ensuring that all staff employed on projects (including research assistants, technicians and clerical staff) act in accordance with the University's ethical principles, the design of the research described in this proposal and any conditions attached to its approval.*

3. Funded Research
Funding body (if any)

School of Psychology, Plymouth University

Is there a potential conflict of interest in the research arising from the source of the funding for the research (for example, a tobacco company funding a study of the effects of smoking on lung function)?

Yes ☐
No ☒

If the answer to the above question is yes, please outline the nature of the potential conflict of interest and how you will address this:

4 Duration of project/programme with dates*:

Immediately – April 2017

**Approval is granted for the duration of projects or for a maximum of three years in the case of programmes. Further approval is necessary for any extension of programmes.*

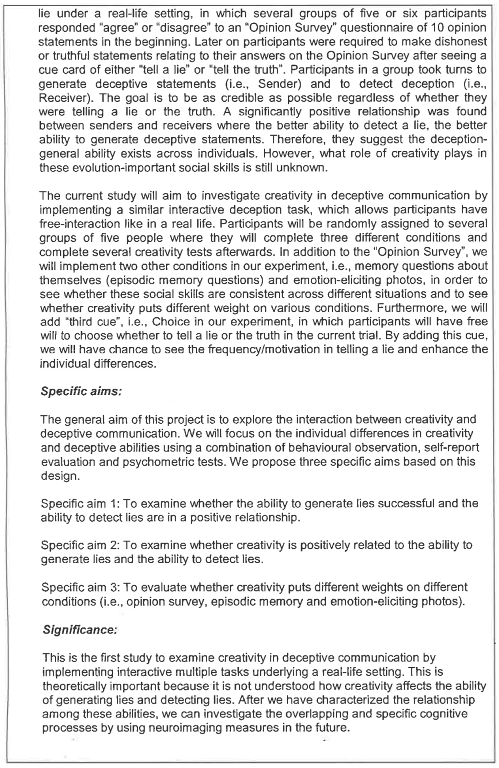
5. Research Outline:

***Background*:**

Is a good judge also a good player? It is common to hear that current good judges were good players in their early ages. In other words, it might require personal experience to know a variety of skills in a specific field to become a good judge. Could the similar concept by applied to generating lies and detecting lies?

One robust finding in the behavioural science has shown that people just are slightly better than chance in detecting lies (Bond & DePaulo, 2006, 2008). On the other hand, fewer studies focus on individual difference in generating lies. Consequently, little research directly investigates the relationship between generating lies and detecting lies. In addition, creativity has been found to be positively related to dishonesty (Gino & Ariely, 2012 & Gino & Wiltermuth, 2014). To our knowledge, no previous study has investigated the role of creativity in detecting lies and whether creativity puts more weight on generating lies or detecting lies. The current research is to explore the relationship between lie-production and lie-detection and what role creativity plays to contribute the individual differences in these social skills.

To our knowledge, to directly investigate the relationship between the ability to lie and the ability to detect lies successfully has only been done by Wright et al. (2012). In that study, a novel competitive interactive deception task (DeceIT) was implemented to measure performance in both the production of lie and detection of



lie under a real-life setting, in which several groups of five or six participants responded "agree" or "disagree" to an "Opinion Survey" questionnaire of 10 opinion statements in the beginning. Later on participants were required to make dishonest or truthful statements relating to their answers on the Opinion Survey after seeing a cue card of either "tell a lie" or "tell the truth". Participants in a group took turns to generate deceptive statements (i.e., Sender) and to detect deception (i.e., Receiver). The goal is to be as credible as possible regardless of whether they were telling a lie or the truth. A significantly positive relationship was found between senders and receivers where the better ability to detect a lie, the better ability to generate deceptive statements. Therefore, they suggest the deception-general ability exists across individuals. However, what role of creativity plays in these evolution-important social skills is still unknown.

The current study will aim to investigate creativity in deceptive communication by implementing a similar interactive deception task, which allows participants have free-interaction like in a real life. Participants will be randomly assigned to several groups of five people where they will complete three different conditions and complete several creativity tests afterwards. In addition to the "Opinion Survey", we will implement two other conditions in our experiment, i.e., memory questions about themselves (episodic memory questions) and emotion-eliciting photos, in order to see whether these social skills are consistent across different situations and to see whether creativity puts different weight on various conditions. Furthermore, we will add "third cue", i.e., Choice in our experiment, in which participants will have free will to choose whether to tell a lie or the truth in the current trial. By adding this cue, we will have chance to see the frequency/motivation in telling a lie and enhance the individual differences.

Specific aims:

The general aim of this project is to explore the interaction between creativity and deceptive communication. We will focus on the individual differences in creativity and deceptive abilities using a combination of behavioural observation, self-report evaluation and psychometric tests. We propose three specific aims based on this design.

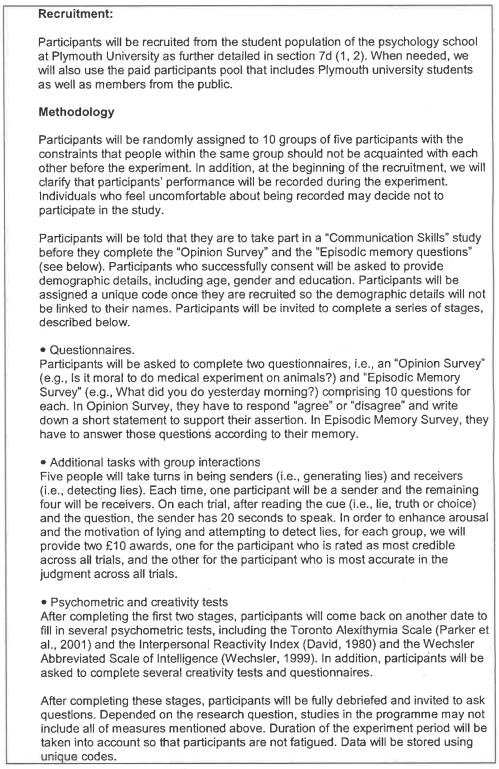
Specific aim 1: To examine whether the ability to generate lies successful and the ability to detect lies are in a positive relationship.

Specific aim 2: To examine whether creativity is positively related to the ability to generate lies and the ability to detect lies.

Specific aim 3: To evaluate whether creativity puts different weights on different conditions (i.e., opinion survey, episodic memory and emotion-eliciting photos).

Significance:

This is the first study to examine creativity in deceptive communication by implementing interactive multiple tasks underlying a real-life setting. This is theoretically important because it is not understood how creativity affects the ability of generating lies and detecting lies. After we have characterized the relationship among these abilities, we can investigate the overlapping and specific cognitive processes by using neuroimaging measures in the future.



Recruitment:

Participants will be recruited from the student population of the psychology school at Plymouth University as further detailed in section 7d (1, 2). When needed, we will also use the paid participants pool that includes Plymouth university students as well as members from the public.

Methodology

Participants will be randomly assigned to 10 groups of five participants with the constraints that people within the same group should not be acquainted with each other before the experiment. In addition, at the beginning of the recruitment, we will clarify that participants' performance will be recorded during the experiment. Individuals who feel uncomfortable about being recorded may decide not to participate in the study.

Participants will be told that they are to take part in a "Communication Skills" study before they complete the "Opinion Survey" and the "Episodic memory questions" (see below). Participants who successfully consent will be asked to provide demographic details, including age, gender and education. Participants will be assigned a unique code once they are recruited so the demographic details will not be linked to their names. Participants will be invited to complete a series of stages, described below.

- Questionnaires.

Participants will be asked to complete two questionnaires, i.e., an "Opinion Survey" (e.g., Is it moral to do medical experiment on animals?) and "Episodic Memory Survey" (e.g., What did you do yesterday morning?) comprising 10 questions for each. In Opinion Survey, they have to respond "agree" or "disagree" and write down a short statement to support their assertion. In Episodic Memory Survey, they have to answer those questions according to their memory.

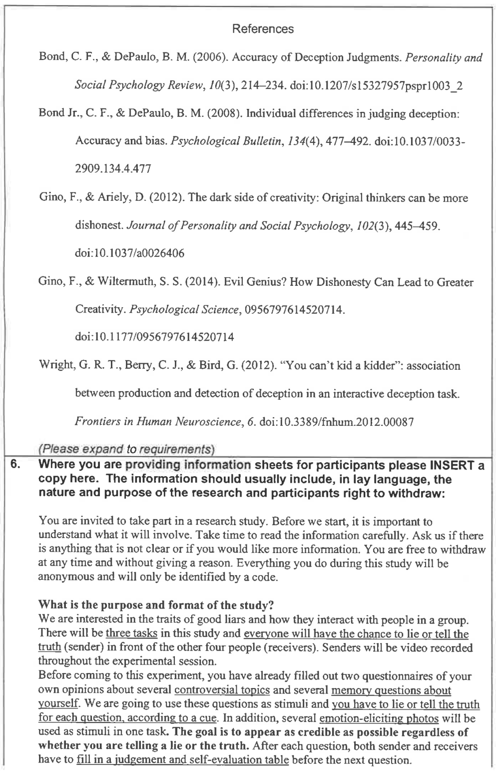
- Additional tasks with group interactions

Five people will take turns in being senders (i.e., generating lies) and receivers (i.e., detecting lies). Each time, one participant will be a sender and the remaining four will be receivers. On each trial, after reading the cue (i.e., lie, truth or choice) and the question, the sender has 20 seconds to speak. In order to enhance arousal and the motivation of lying and attempting to detect lies, for each group, we will provide two £10 awards, one for the participant who is rated as most credible across all trials, and the other for the participant who is most accurate in the judgment across all trials.

- Psychometric and creativity tests

After completing the first two stages, participants will come back on another date to fill in several psychometric tests, including the Toronto Alexithymia Scale (Parker et al., 2001) and the Interpersonal Reactivity Index (David, 1980) and the Wechsler Abbreviated Scale of Intelligence (Wechsler, 1999). In addition, participants will be asked to complete several creativity tests and questionnaires.

After completing these stages, participants will be fully debriefed and invited to ask questions. Depended on the research question, studies in the programme may not include all of measures mentioned above. Duration of the experiment period will be taken into account so that participants are not fatigued. Data will be stored using unique codes.



References

Bond, C. F., & DePaulo, B. M. (2006). Accuracy of Deception Judgments. *Personality and Social Psychology Review*, *10*(3), 214–234. doi:10.1207/s15327957pspr1003_2

Bond Jr., C. F., & DePaulo, B. M. (2008). Individual differences in judging deception: Accuracy and bias. *Psychological Bulletin*, *134*(4), 477–492. doi:10.1037/0033-2909.134.4.477

Gino, F., & Ariely, D. (2012). The dark side of creativity: Original thinkers can be more dishonest. *Journal of Personality and Social Psychology*, *102*(3), 445–459. doi:10.1037/a0026406

Gino, F., & Wiltermuth, S. S. (2014). Evil Genius? How Dishonesty Can Lead to Greater Creativity. *Psychological Science*, 0956797614520714. doi:10.1177/0956797614520714

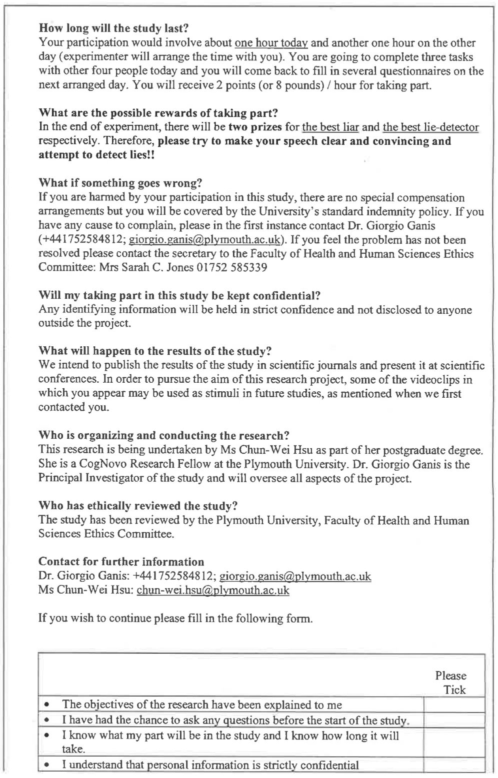
Wright, G. R. T., Berry, C. J., & Bird, G. (2012). "You can't kid a kidder": association between production and detection of deception in an interactive deception task. *Frontiers in Human Neuroscience*, *6*. doi:10.3389/fnhum.2012.00087

(Please expand to requirements)

6. Where you are providing information sheets for participants please INSERT a copy here. The information should usually include, in lay language, the nature and purpose of the research and participants right to withdraw:

You are invited to take part in a research study. Before we start, it is important to understand what it will involve. Take time to read the information carefully. Ask us if there is anything that is not clear or if you would like more information. You are free to withdraw at any time and without giving a reason. Everything you do during this study will be anonymous and will only be identified by a code.

What is the purpose and format of the study?
We are interested in the traits of good liars and how they interact with people in a group. There will be three tasks in this study and everyone will have the chance to lie or tell the truth (sender) in front of the other four people (receivers). Senders will be video recorded throughout the experimental session.
Before coming to this experiment, you have already filled out two questionnaires of your own opinions about several controversial topics and several memory questions about yourself. We are going to use these questions as stimuli and you have to lie or tell the truth for each question, according to a cue. In addition, several emotion-eliciting photos will be used as stimuli in one task. **The goal is to appear as credible as possible regardless of whether you are telling a lie or the truth.** After each question, both sender and receivers have to fill in a judgement and self-evaluation table before the next question.



How long will the study last?
Your participation would involve about one hour today and another one hour on the other day (experimenter will arrange the time with you). You are going to complete three tasks with other four people today and you will come back to fill in several questionnaires on the next arranged day. You will receive 2 points (or 8 pounds) / hour for taking part.

What are the possible rewards of taking part?
In the end of experiment, there will be **two prizes** for the best liar and the best lie-detector respectively. Therefore, **please try to make your speech clear and convincing and attempt to detect lies!!**

What if something goes wrong?
If you are harmed by your participation in this study, there are no special compensation arrangements but you will be covered by the University's standard indemnity policy. If you have any cause to complain, please in the first instance contact Dr. Giorgio Ganis (+441752584812; giorgio.ganis@plymouth.ac.uk). If you feel the problem has not been resolved please contact the secretary to the Faculty of Health and Human Sciences Ethics Committee: Mrs Sarah C. Jones 01752 585339

Will my taking part in this study be kept confidential?
Any identifying information will be held in strict confidence and not disclosed to anyone outside the project.

What will happen to the results of the study?
We intend to publish the results of the study in scientific journals and present it at scientific conferences. In order to pursue the aim of this research project, some of the videoclips in which you appear may be used as stimuli in future studies, as mentioned when we first contacted you.

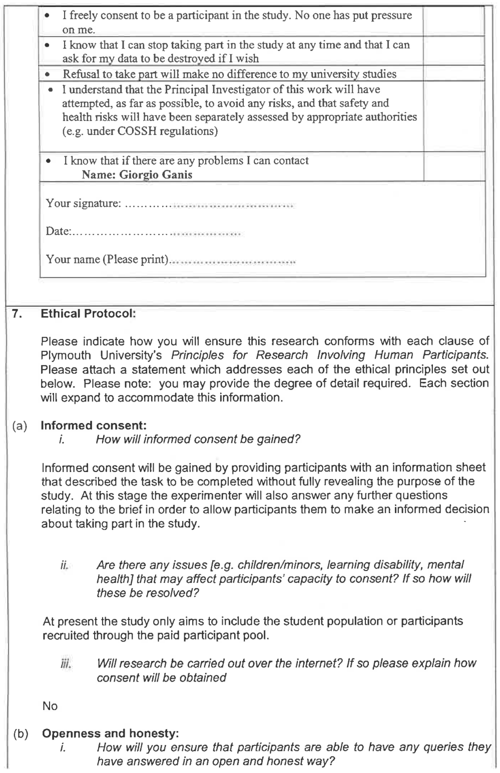
Who is organizing and conducting the research?
This research is being undertaken by Ms Chun-Wei Hsu as part of her postgraduate degree. She is a CogNovo Research Fellow at the Plymouth University. Dr. Giorgio Ganis is the Principal Investigator of the study and will oversee all aspects of the project.

Who has ethically reviewed the study?
The study has been reviewed by the Plymouth University, Faculty of Health and Human Sciences Ethics Committee.

Contact for further information
Dr. Giorgio Ganis: +441752584812; giorgio.ganis@plymouth.ac.uk
Ms Chun-Wei Hsu: chun-wei.hsu@plymouth.ac.uk

If you wish to continue please fill in the following form.

	Please Tick
• The objectives of the research have been explained to me	
• I have had the chance to ask any questions before the start of the study.	
• I know what my part will be in the study and I know how long it will take.	
• I understand that personal information is strictly confidential	



• I freely consent to be a participant in the study. No one has put pressure on me.	
• I know that I can stop taking part in the study at any time and that I can ask for my data to be destroyed if I wish	
• Refusal to take part will make no difference to my university studies	
• I understand that the Principal Investigator of this work will have attempted, as far as possible, to avoid any risks, and that safety and health risks will have been separately assessed by appropriate authorities (e.g. under COSSH regulations)	
• I know that if there are any problems I can contact **Name: Giorgio Ganis**	

Your signature: ……………………………………

Date:……………………………………

Your name (Please print)…………………………

7. Ethical Protocol:

Please indicate how you will ensure this research conforms with each clause of Plymouth University's *Principles for Research Involving Human Participants*. Please attach a statement which addresses each of the ethical principles set out below. Please note: you may provide the degree of detail required. Each section will expand to accommodate this information.

(a) **Informed consent:**

i. How will informed consent be gained?

Informed consent will be gained by providing participants with an information sheet that described the task to be completed without fully revealing the purpose of the study. At this stage the experimenter will also answer any further questions relating to the brief in order to allow participants them to make an informed decision about taking part in the study.

ii. Are there any issues [e.g. children/minors, learning disability, mental health] that may affect participants' capacity to consent? If so how will these be resolved?

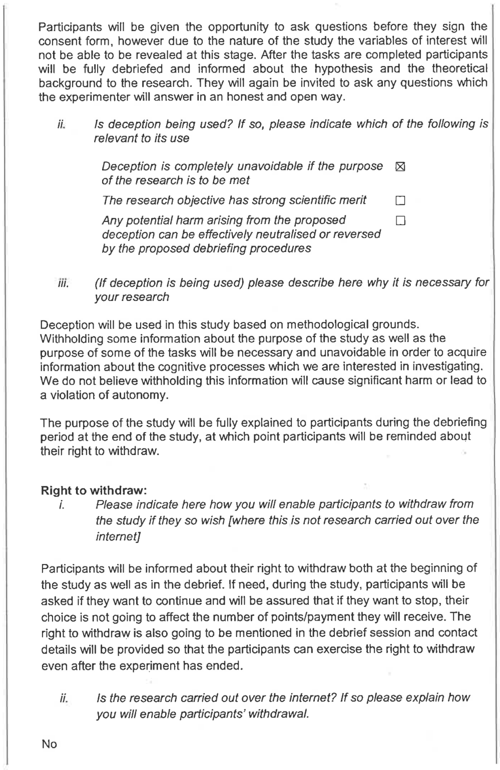
At present the study only aims to include the student population or participants recruited through the paid participant pool.

iii. Will research be carried out over the internet? If so please explain how consent will be obtained

No

(b) **Openness and honesty:**

i. How will you ensure that participants are able to have any queries they have answered in an open and honest way?



Participants will be given the opportunity to ask questions before they sign the consent form, however due to the nature of the study the variables of interest will not be able to be revealed at this stage. After the tasks are completed participants will be fully debriefed and informed about the hypothesis and the theoretical background to the research. They will again be invited to ask any questions which the experimenter will answer in an honest and open way.

ii. Is deception being used? If so, please indicate which of the following is relevant to its use

Deception is completely unavoidable if the purpose of the research is to be met ☒
The research objective has strong scientific merit ☐
Any potential harm arising from the proposed deception can be effectively neutralised or reversed by the proposed debriefing procedures ☐

iii. (If deception is being used) please describe here why it is necessary for your research

Deception will be used in this study based on methodological grounds. Withholding some information about the purpose of the study as well as the purpose of some of the tasks will be necessary and unavoidable in order to acquire information about the cognitive processes which we are interested in investigating. We do not believe withholding this information will cause significant harm or lead to a violation of autonomy.

The purpose of the study will be fully explained to participants during the debriefing period at the end of the study, at which point participants will be reminded about their right to withdraw.

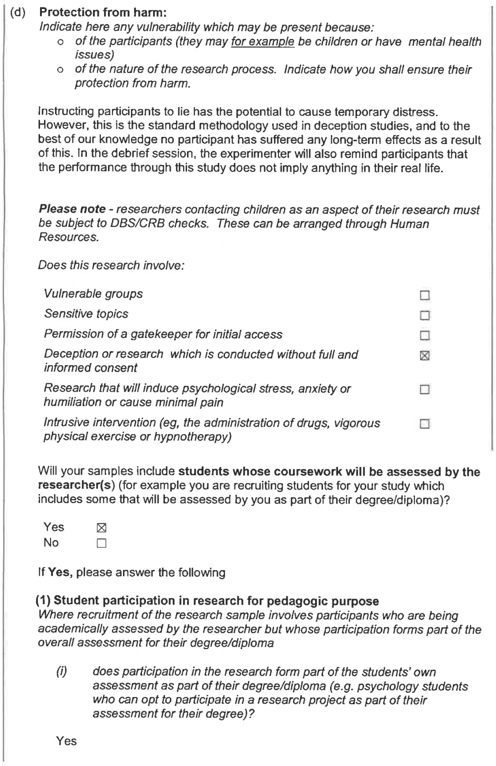
Right to withdraw:

i. Please indicate here how you will enable participants to withdraw from the study if they so wish [where this is not research carried out over the internet]

Participants will be informed about their right to withdraw both at the beginning of the study as well as in the debrief. If need, during the study, participants will be asked if they want to continue and will be assured that if they want to stop, their choice is not going to affect the number of points/payment they will receive. The right to withdraw is also going to be mentioned in the debrief session and contact details will be provided so that the participants can exercise the right to withdraw even after the experiment has ended.

ii. Is the research carried out over the internet? If so please explain how you will enable participants' withdrawal.

No



(d) **Protection from harm:**
Indicate here any vulnerability which may be present because:

- *of the participants (they may for example be children or have mental health issues)*
- *of the nature of the research process. Indicate how you shall ensure their protection from harm.*

Instructing participants to lie has the potential to cause temporary distress. However, this is the standard methodology used in deception studies, and to the best of our knowledge no participant has suffered any long-term effects as a result of this. In the debrief session, the experimenter will also remind participants that the performance through this study does not imply anything in their real life.

***Please note** - researchers contacting children as an aspect of their research must be subject to DBS/CRB checks. These can be arranged through Human Resources.*

Does this research involve:

Vulnerable groups	☐
Sensitive topics	☐
Permission of a gatekeeper for initial access	☐
Deception or research which is conducted without full and informed consent	☒
Research that will induce psychological stress, anxiety or humiliation or cause minimal pain	☐
Intrusive intervention (eg, the administration of drugs, vigorous physical exercise or hypnotherapy)	☐

Will your samples include **students whose coursework will be assessed by the researcher(s)** (for example you are recruiting students for your study which includes some that will be assessed by you as part of their degree/diploma)?

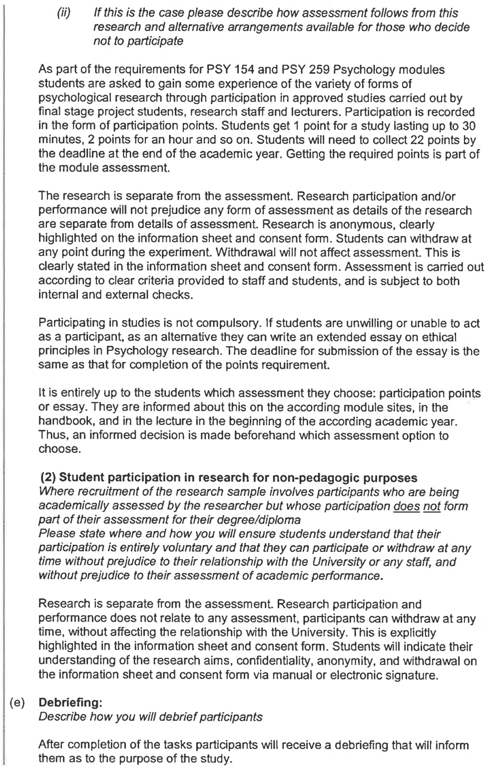
Yes ☒
No ☐

If **Yes,** please answer the following

(1) Student participation in research for pedagogic purpose
Where recruitment of the research sample involves participants who are being academically assessed by the researcher but whose participation forms part of the overall assessment for their degree/diploma

(i) *does participation in the research form part of the students' own assessment as part of their degree/diploma (e.g. psychology students who can opt to participate in a research project as part of their assessment for their degree)?*

Yes



(ii) *If this is the case please describe how assessment follows from this research and alternative arrangements available for those who decide not to participate*

As part of the requirements for PSY 154 and PSY 259 Psychology modules students are asked to gain some experience of the variety of forms of psychological research through participation in approved studies carried out by final stage project students, research staff and lecturers. Participation is recorded in the form of participation points. Students get 1 point for a study lasting up to 30 minutes, 2 points for an hour and so on. Students will need to collect 22 points by the deadline at the end of the academic year. Getting the required points is part of the module assessment.

The research is separate from the assessment. Research participation and/or performance will not prejudice any form of assessment as details of the research are separate from details of assessment. Research is anonymous, clearly highlighted on the information sheet and consent form. Students can withdraw at any point during the experiment. Withdrawal will not affect assessment. This is clearly stated in the information sheet and consent form. Assessment is carried out according to clear criteria provided to staff and students, and is subject to both internal and external checks.

Participating in studies is not compulsory. If students are unwilling or unable to act as a participant, as an alternative they can write an extended essay on ethical principles in Psychology research. The deadline for submission of the essay is the same as that for completion of the points requirement.

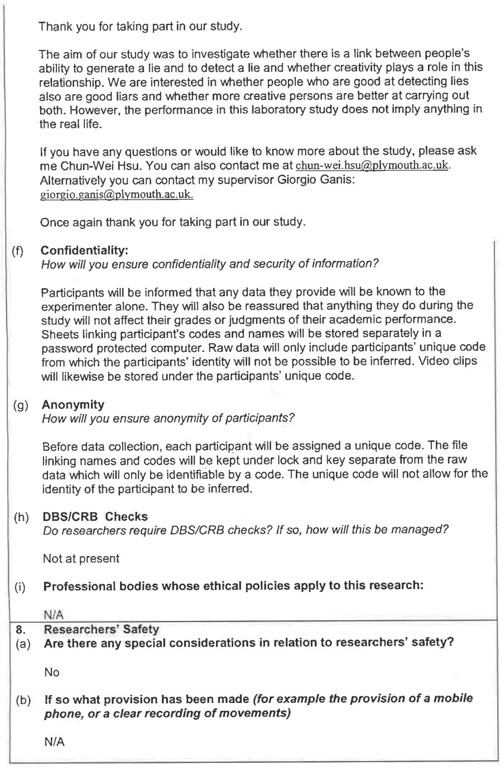
It is entirely up to the students which assessment they choose: participation points or essay. They are informed about this on the according module sites, in the handbook, and in the lecture in the beginning of the according academic year. Thus, an informed decision is made beforehand which assessment option to choose.

(2) Student participation in research for non-pedagogic purposes
Where recruitment of the research sample involves participants who are being academically assessed by the researcher but whose participation does not form part of their assessment for their degree/diploma
Please state where and how you will ensure students understand that their participation is entirely voluntary and that they can participate or withdraw at any time without prejudice to their relationship with the University or any staff, and without prejudice to their assessment of academic performance.

Research is separate from the assessment. Research participation and performance does not relate to any assessment, participants can withdraw at any time, without affecting the relationship with the University. This is explicitly highlighted in the information sheet and consent form. Students will indicate their understanding of the research aims, confidentiality, anonymity, and withdrawal on the information sheet and consent form via manual or electronic signature.

(e) **Debriefing:**
Describe how you will debrief participants

After completion of the tasks participants will receive a debriefing that will inform them as to the purpose of the study.



Thank you for taking part in our study.

The aim of our study was to investigate whether there is a link between people's ability to generate a lie and to detect a lie and whether creativity plays a role in this relationship. We are interested in whether people who are good at detecting lies also are good liars and whether more creative persons are better at carrying out both. However, the performance in this laboratory study does not imply anything in the real life.

If you have any questions or would like to know more about the study, please ask me Chun-Wei Hsu. You can also contact me at chun-wei.hsu@plymouth.ac.uk. Alternatively you can contact my supervisor Giorgio Ganis: giorgio.ganis@plymouth.ac.uk.

Once again thank you for taking part in our study.

(f) **Confidentiality:**
How will you ensure confidentiality and security of information?

Participants will be informed that any data they provide will be known to the experimenter alone. They will also be reassured that anything they do during the study will not affect their grades or judgments of their academic performance. Sheets linking participant's codes and names will be stored separately in a password protected computer. Raw data will only include participants' unique code from which the participants' identity will not be possible to be inferred. Video clips will likewise be stored under the participants' unique code.

(g) **Anonymity**
How will you ensure anonymity of participants?

Before data collection, each participant will be assigned a unique code. The file linking names and codes will be kept under lock and key separate from the raw data which will only be identifiable by a code. The unique code will not allow for the identity of the participant to be inferred.

(h) **DBS/CRB Checks**
Do researchers require DBS/CRB checks? If so, how will this be managed?

Not at present

(i) **Professional bodies whose ethical policies apply to this research:**

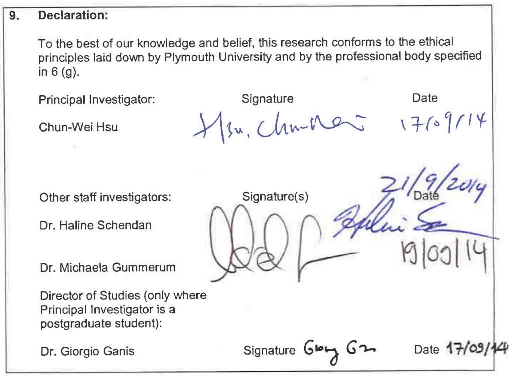
N/A

8. **Researchers' Safety**
(a) **Are there any special considerations in relation to researchers' safety?**

No

(b) **If so what provision has been made *(for example the provision of a mobile phone, or a clear recording of movements)***

N/A



9. **Declaration:**

To the best of our knowledge and belief, this research conforms to the ethical principles laid down by Plymouth University and by the professional body specified in 6 (g).

Principal Investigator:	Signature	Date
Chun-Wei Hsu	Hsu, Chun-Wei	17/09/14
Other staff investigators:	Signature(s)	Date
Dr. Haline Schendan		21/9/2014
Dr. Michaela Gummerum		19/09/14
Director of Studies (only where Principal Investigator is a postgraduate student):		
Dr. Giorgio Ganis	Signature	Date 17/09/14

APPENDIX 2A

Subject information and consent form for Study 1

You are invited to take part in a research study. Before we start, it is important to understand what it will involve. Take time to read the information carefully. Ask us if there is anything that is not clear or if you would like more information. You are free to withdraw at any time and without giving a reason. Everything you do during this study will be anonymous and will only be identified by a code.

What is the purpose and format of the study?

We are interested in the traits of good liars and how they interact with people in a group. There will be three tasks in this study and everyone will have the chance to lie or tell the truth (sender) in front of the other four people (receivers). Senders will be video recorded throughout the experimental session.

Before this experiment starts, you have already filled out two questionnaires of your own opinions about several controversial topics and several memory questions about yourself. We are going to use these questions as stimuli and you have to lie or tell the truth for each question, according to a cue. In addition, several emotion-eliciting photos will be used as stimuli in one task. **The goal is to appear as credible as possible regardless of whether you are telling a lie or the truth.** After each question, both sender and receivers have to fill in a judgement and self-evaluation table before the next question.

How long will the study last?

Your participation would involve about two hours today and another one hour on the other day (experimenter will arrange the time with you). You are going to complete three tasks with other four people today and you will come back to fill in several questionnaires on the next arranged day. You will receive 2 points (or 8 pounds) / hour for taking part.

What are the possible rewards of taking part?

In the end of experiment, we will prepare **six PRIZES** (Amazon coupons) for **the best three liars and the best three lie-detectors**. (£30 for the first place; £20 for the second place; £10 for the third place) Therefore, **please try to make your speech clear and convincing and attempt to detect lies!!**

What if something goes wrong?

If you are harmed by your participation in this study, there are no special compensation arrangements but you will be covered by the University's standard indemnity policy. If you have any cause to complain, please in the first instance contact Dr. Giorgio Ganis (+441752584812; giorgio.ganis@plymouth.ac.uk). If you feel the problem has not been resolved please contact the secretary to the Faculty of Health and Human Sciences Ethics Committee: Mrs Sarah C. Jones 01752 585339

Will my taking part in this study be kept confidential?

Any identifying information will be held in strict confidence and not disclosed to anyone outside the project.

What will happen to the results of the study?

We intend to publish the results of the study in scientific journals and present it at scientific conferences. In order to pursue the aim of this research project, some of the videoclips in which you appear may be used as stimuli in future studies, as mentioned when we first contacted you.

Who is organizing and conducting the research?

This research is being undertaken by Ms Chun-Wei Hsu as part of her postgraduate degree. She is a CogNovo Research Fellow at the Plymouth University. Dr. Giorgio Ganis is the Principal Investigator of the study and will oversee all aspects of the project.

Who has ethically reviewed the study?

The study has been reviewed by the Plymouth University, Faculty of Health and Human Sciences Ethics Committee.

Contact for further information

Dr. Giorgio Ganis: +441752584812; giorgio.ganis@plymouth.ac.uk
Ms Chun-Wei Hsu: chun-wei.hsu@plymouth.ac.uk

If you wish to continue please fill in the following form.

	Please Tick
• The objectives of the research have been explained to me	
• I have had the chance to ask any questions before the start of the study.	
• I know what my part will be in the study and I know how long it will take.	
• I understand that personal information is strictly confidential	
• I freely consent to be a participant in the study. No one has put pressure on me.	
• I know that I can stop taking part in the study at any time and that I can ask for my data to be destroyed if I wish	
• Refusal to take part will make no difference to my university studies	
• I understand that the Principal Investigator of this work will have attempted, as far as possible, to avoid any risks, and that safety and health risks will have been separately assessed by appropriate authorities (e.g. under COSSH regulations)	
• I know that if there are any problems I can contact **Name: Giorgio Ganis**	

Your signature: ……………………………………

Date:……………………………………

Your name (Please print)…………………………..

APPENDIX 2B

Opinion survey and episodic memory items used in Study 1

Opinion survey

Practice:

1. Do you support immigration?
2. Do you support legalizing prostitution?

Real task:

1. Should we stop doing medical experiments on animals?
2. Should GM food be banned?
3. Should people have right to die?
4. Should cosmetic surgery be covered by NHS?
5. Should obese people pay more for plane tickets?
6. Is it ethical to take pills to make people smarter?
7. Should smokers pay more for health services?
8. Is it ethical to have a child outside of marriage?
9. Should condoms be distributed freely in high schools?
10. Does God exist?

Episodic memory

Practice:

1. Please describe the best trip you have ever had.
2. Please describe your best friend.

Real task:

1. What did you do last weekend?
2. How did you spend your last birthday?
3. Where did you travel during your summer vacation?
4. What did you do last night?
5. Please describe your hometown?
6. How did you spend your last Christmas?
7. Please describe the most recent restaurant you ate at.
8. What did you do yesterday morning?
9. Please describe the most recent movie you saw.
10. How did you spend your New Year’s Eve?

APPENDIX 3A

Subject information and consent form for Study 2

You are invited to take part in a research study. Before we start, it is important to understand what it will involve. Take time to read the information carefully. Ask us if there is anything that is not clear or if you would like more information. You are free to withdraw at any time and without giving a reason. Everything you do during this study will be anonymous and will only be identified by a code.

What is the purpose and format of the study?
There will be three games in the study. Participants will wear on an electroencephalogram (EEG) cap and brain waves will be recorded during the experiment.

The first game is a simple observation game. You are going to be shown a series of video clips and press "Yes" button when you see the target/behaviour, and otherwise press "No" button.
The second game is loosely based on the popular card game of cheat. The objective is for the participants to deceive the judge in order to win as much money as possible. How much money they have won for the trial and total amount of money will be presented at the end of each trial. Participants will be awarded this amount of money **in cash**. **The goal is to appear as credible as possible regardless of whether you are telling a lie or the truth**.
Participants will be filmed for the duration of the task —a live stream will be presented to the Judge.
The final game is for participants to become a Judge and to detect whether the player in the clips is lying or not. In each trial to participants will be presented with a short clip, in which the player in the clip is playing the same bluffing game as you just played. The mission is to judge the player is telling a lie or the truth. **Please try your best to detect lies in the player!**

How long will the study last?
Your participation would involve about 2 hours today and another 1.5 hours on the other day.
You will receive 2 points (or 8 pounds) / hour for taking part.

What are the rewards of taking part?
At the end of the experiment, participants will be presented with the final amount they have won and will then be awarded this amount **in cash** alongside payment for participating in the experiment.

What if something goes wrong?
If you are harmed by your participation in this study, there are no special compensation arrangements but you will be covered by the University's standard indemnity policy. If you have any cause to complain, please in the first instance contact Dr. Giorgio Ganis (+441752584812; giorgio.ganis@plymouth.ac.uk). If you feel the problem has not been resolved please contact the secretary to the Faculty of Health and Human Sciences Ethics Committee: Mrs Sarah C. Jones 01752 585339
Will my taking part in this study be kept confidential?
Any identifying information will be held in strict confidence and not disclosed to anyone outside the project.

What will happen to the results of the study?
We intend to publish the results of the study in scientific journals and present it at scientific conferences. In order to pursue the aim of this research project, some of the videoclips in

which you appear may be used as stimuli in future studies, as mentioned when we first contacted you.

Who is organizing and conducting the research?

This research is being undertaken by Ms Chun-Wei Hsu as part of her postgraduate degree. She is a CogNovo Research Fellow at the Plymouth University. Dr. Giorgio Ganis is the Principal Investigator of the study and will oversee all aspects of the project.

Who has ethically reviewed the study?

The study has been reviewed by the Plymouth University, Faculty of Health and Human Sciences Ethics Committee.

Contact for further information

Dr. Giorgio Ganis: +441752584812; giorgio.ganis@plymouth.ac.uk
Ms Chun-Wei Hsu: chun-wei.hsu@plymouth.ac.uk
Ms Sarah Hounsell: sarah.hounsell@plymouth.ac.uk

If you wish to continue please fill in the following form.

	Please Tick
• The objectives of the research have been explained to me	
• I have had the chance to ask any questions before the start of the study.	
• I know what my part will be in the study and I know how long it will take.	
• I understand that personal information is strictly confidential	
• I freely consent to be a participant in the study. No one has put pressure on me.	
• I know that I can stop taking part in the study at any time and that I can ask for my data to be destroyed if I wish	
• Refusal to take part will make no difference to my university studies	
• I understand that the Principal Investigator of this work will have attempted, as far as possible, to avoid any risks, and that safety and health risks will have been separately assessed by appropriate authorities (e.g. under COSSH regulations)	
• I know that if there are any problems I can contact **Name: Giorgio Ganis or Chun-Wei Hsu**	

Your signature: ……………………………………

Date:……………………………………

Your name (Please print)…………………………..

APPENDIX 3B

Follow-up questions for Study 2

Subject no.:

FOLLOW-UP QUESTIONS

Thank you for taking part in the first session of our study.

Please complete the following two questions related to the experiment you just did.

For bluffing game:

- Did you find any difficulties in understanding the rules of the game? If yes, please describe it

____

____

- Please write down any strategies you used to tell convincing lies.

____

____

- Did you find the bluffing game interesting? Could you concentrate on the game all the time through the experiment?

____

____

For detection game:

- Please write down all strategies you used to judge if the player was lying.

____

____

- Did you find the detection game interesting? Could you concentrate on the game all the time through the experiment?

____

Once again thank you for taking part in our study.

APPENDIX 3C

Supplementary results for Study 2

1. Method

1.1 *Bluffing game*

For the bluffing game, Repeated-measures of ANOVAs were conducted on the mean amplitude of the average ERPs of interest for the comparison between Cued truths and Small lies and the comparison between Uncued truths and Small lies, respectively. For the N200 and the P300, "midline" ANOVAs were carried out on midline electrodes using 2 factors: Response (Cued/Uncued truth, Small lie) and Site (Fz, Cz, Pz and Oz). To assess the overall pattern of results, "lateral" ANOVAs were also carried out on lateral electrodes (14 pairs) using three factors: Response, Site, and Hemisphere. For the MFN and the PRP, ANOVAs were carried out using 2 factors: Response (Cued/Uncued truth, Small lie) and Site (Fz, F3, F4, FC1, FC2 and Cz). For the LPC, the ANOVAs were carried out using 2 factors: Response (Cued/Uncued truth, Small lie) and Site (Pz, P3, P4, CP1 and CP2). A diagram showing the sites analysed for the MFN, PRP and LPC components is shown in Figure 3-3. For the BP, ANOVAs were carried out using 3 factors: Response (Cued/Uncued truth, Small lie), Time window (-800 to -600 ms, -600 to -400 ms) and Site (Fz, Cz).

Another comparison between Cued truth and Uncued truth conditions was also conducted to examine whether Uncued truths would require additional cognitive effort compared to Cued truths, even though both involved truth-telling. This comparison could also provide further evidence about whether differences in the ERPs of interest were driven by deception processes or more general decision-making processes.

Since 20 participants (out of 36) had less than 10 valid trials in the Big lie condition, the current ERP results only included Small lies in analyses. Nevertheless, we also carried out a weighted average to combine Small lie and Big lie into a "Total lie" category and performed the same analyses as above. These results are only shown in the tables.

1.2 *Lie detection*

Repeated-measures ANOVAs were carried out on these components using 3 factors: Group (good, bad), Stimulus (truth, lie) and Judgement (truth, lie). In addition to the analyses of specific visual ERPs, the overall pattern of results was explored using time windows every 100 ms between 0 and 900 ms at five different regions of the scalp: (i) The midline region included four electrode sites (Fz, Cz, Pz and Oz); (ii) the lateral regions included frontal (Fp1/Fp2, AF3/AF4, F7/F8, F3/F4), central (FC1/FC2, C3/C4, CP1/CP2), temporal (FC5/FC6, T7/T8, CP5/CP6) and posterior (P3.P4, P7/P8, PO3/PO4, O1/O2) sites. The analysis of the midline region was conducted separately from the four lateral regions. Repeated-measures ANOVAs were carried out using 4 factors: Group (good, bad), Regions (frontal, central, temporal, posterior), Stimulus (truth, lie) and Judgement (truth, lie). The analysis of the midline region only included the factors of Group, Stimulus and Judgement.

In addition to the onset of the clip, we also locked to the onset of the speaking cue (i.e., audio onset). We performed the same regional analyses on the audio onset to demonstrate the overall pattern of ERP waveforms.

Because of the risk of false-positive effects in the multiple comparisons in the regional analyses, results were considered significant only if there were at least two successive time bins significant in a given region (effects lasting more than ≥ 200 ms) (Nobre, Rao, & Chelazzi, 2006).

1.3 *Lie detection and Head movement*

To assess potential ERP differences between judging if a person was lying or telling the truth and judging a physical feature (i.e., head motion), the same analyses were carried out as in the lie detection task but using the factor Task (detection, head movement), instead of Group.

All *p* values for ERP results were adjusted with the Greenhouse-Geisser epsilon correction (reported if smaller than 1) for nonsphericity when necessary. Effect sizes (partial eta squared [η^2]) are reported, with 0.01, 0.06 and 0.14 considered small, medium, and large effect sizes, respectively. Cohen's *d* was calculated as effect size for paired data, with 0.20, 0.50 and 0.80 considered small, medium and large effect sizes. For all analyses, only significant effects are reported, unless otherwise specified. Finally, trends at $.05 < p < .1$ are reported when relevant. All statistical analyses were carried out using IBM SPSS Statistics (Version 23).

2. Results

2.1 *Bluffing game*

2.1.1 *N200* (250-350 ms, Stimulus-locked)

The aggregated grand average waveform from trials revealed an N200 around 300 ms after the stimulus at fronto-central electrodes (Figure 3-4b). The two midline ANOVAs (Cued truth vs. Small lie and Uncued truth vs. Small lie) showed no main effect of response and no interaction between response and site. The lateral ANOVAs also revealed neither a main effect of response nor an interaction between response and site (Table 3C-A1). There was an interaction between response and hemisphere when comparing Cued truths and Small lies. Planned focal analyses showed a larger N200 on the left than right hemisphere in both Cued truth and Small lie conditions, and the difference between hemispheres was bigger in the Cued truth condition. Since we focused on the frontal N200, this interaction between response and hemisphere will not be further discussed.

Additional the midline ANOVA comparing Cued truths and Uncued truths showed neither a main effect of response nor an interaction. The lateral ANOVA revealed an interaction between response and site, but not a main effect of response (Table 3C-A1). To further examine the interaction, Uncued truths showed a trend for a larger N200 compared with Cued truths at most of sites, but a few occipito-parietal sites showed an opposite trend with a smaller N200 for Uncued truths compared with Cued truths. Since planned focal analyses showed no site with a significant difference between Cued truths and Uncued truths, this interaction will not be further discussed.

In summary, these results showed no frontal N200 effects for lying compared with Cued or Uncued truths. Also, no frontal N200 differences were found between Cued truths and Uncued truths.

2.1.2 *P300* (450-650 ms, Stimulus-locked)

The aggregated grand average waveform from trials revealed a P300 around 550 ms after the stimulus at parietal electrodes (Figure 3-4c). For the first midline ANOVA comparing Cued truth and Small lie conditions, a main effect of response showed that lying was associated with a smaller P300 compared with the cued truth. The interaction between response and site was also significant. Planned focal analyses showed a decreased P300 for lying compared with cued truth at Pz (*t*[35] = 4.76, *p* < .001, *d* = 0.79, 95% CI [1.51, 3.77]), and Cz (*t*[35] = 3.45, *p* < 0.01, *d* = 0.58, 95% CI [0.77, 2.97]. In contrast, the midline ANOVA comparing Uncued truths and Small lies showed neither a main effect of response nor an interaction between response and site (Table 3C-A1).

Similarly, the lateral ANOVAs revealed a main effect of response and an interaction between response and site only in the comparison between Cued truth and Small lie conditions (Table 3C-A1).

An additional midline ANOVA comparing Cued truths and Uncued truths showed a main effect of response and an interaction between response and site. Planned focal analyses showed a decreased P300 for Uncued truths compared with Cued truths at Pz (*t*[35] = 5.17, *p* < 0.001, *d* = 0.86, 95% CI [1.49, 3.41]), Cz (*t*[35] = 4.48, *p* < 0.001, *d* = 0.75, 95% CI [1.09, 2.89]) and Fz (*t*[35] = 2.44, *p* < 0.05, *d* = 0.41, 95% CI [0.22, 2.42]). The lateral ANOVA also revealed a main effect of response where Uncued truths were associated with a smaller P300. The interaction between response and site was also significant (Table 3C-A1).

In summary, a smaller P300 for lying was found compared with Cued truths but not Uncued truths. In addition, the P300 for Uncued truths was smaller compared to that for cued truth.

A weighted average was carried out to combine Small and Big lies into a “Total lie” category. The comparison between “Truth” and “Total lie” is shown in Table 3C-A2, and the results are comparable to those in the Small lie condition.

2.1.3 *MFN* (0, 100 ms, Response-locked)

The aggregated grand average waveform from trials revealed that the MFN peaked around 50 ms after the response at fronto-central electrodes (Figure 3-5c). The first ANOVA comparing Cued truths and Small lies revealed an interaction between response and site but no main effect of response (Table 3C-B1). Planned focal analyses showed a larger MFN for Small lies compared to Cued truths at FC1 (*t*[35] = 2.16, *p* < 0.05, *d* = 0.36, 95% CI [0.10, 3.31]), FC2 (*t*[35] = 2.58, *p* < 0.05, *d* = 0.43, 95% CI [0.50, 4.19]) and Cz (*t*[35] = 2.39, *p* < 0.05, *d* = 0.40, 95% CI [0.32, 3.90]). The second ANOVA comparing Uncued truths and Small lies revealed neither a main effect of response nor an interaction between response and site (Table 3C-B1).

An additional ANOVA comparing Cued and Uncued truths showed neither a main effect of response nor an interaction between response and site (Table 3C-B1).

In summary, a larger MFN for lying was found compared with the Cued truths but not with the Uncued truths. In addition, no MFN difference was found between Cued truths and Uncued truths.

2.1.4 *PRP* (-200, 0 ms, Response-locked)

The aggregated grand average waveform from trials revealed the PRP around 200 ms before the response at fronto-central electrodes (Figure 3-5c). The first ANOVA comparing the Cued truth and Small lie conditions revealed an interaction between response and site but no main effect of response (Table 3C-B1). Planned focal analyses showed a smaller PRP for lying compared to cued truth at FC1 (*t*[35] = 2.29, *p* < 0.05, *d* = 0.38, 95% CI [0.16, 2.72]), FC2 (*t*[35] = 3.03, *p* < .01, *d* = 0.51, 95% CI [0.73, 3.68]) and Cz ($t[35] = 3.21, p < .01, d = 0.54$, 95% CI [0.87, 3.87]). Similarly, the second ANOVA comparing the Uncued truth and Small lie condition revealed an interaction between response and site but no main effect of response (Table 3C-B1). However, planned focal analyses showed no site with significant difference between lying and uncued truth.

An additional ANOVA comparing the Cued truth and Uncued truth conditions showed a marginal main effect of response and an interaction between response and site (Table 3C-B1). Planned focal analyses showed a smaller PRP for uncued truth compared with cued truth at FC1 ($t[35] = 2.35, p < 0.05, d = 0.39$, 95% CI [0.21, 2.94]), FC2 ($t[35] = 2.46, p < 0.05, d = 0.41$, 95% CI [0.34, 3.53]) and Cz ($t[35] = 2.28, p < 0.05, d = 0.38$, 95% CI [0.19, 3.22]).

In summary, a smaller PRP for lying was found compared with Cued truths but not Uncued truths. In addition, Uncued truths showed a smaller PRP compared with Cued truths.

2.1.5 *LPC* (-250, 0 ms, Response-locked)

The aggregated grand average waveform from trials revealed the LPC around 250 ms before the response at central-parietal electrodes (Figure 3-5d). For the first ANOVA comparing the Cued truth and Small lie condition, showed a main effect of response indicating that lying reduced LPC amplitude compared to cued truth, as well as an interaction between response and site (Table 3C-B1). Planned focal analyses showed a smaller LPC for lying compared with cued truth at Pz ($t[35] = 4.38, p < .001, d = 0.73$, 95% CI [1.80, 4.92]), P3 ($t[35] = 2.67, p < 0.05, d = 0.45$, 95% CI [0.49, 3.59], P4 ($t[35] = 3.52, p < .01, d = 0.59$, 95% CI [1.13, 4.20], CP1 ($t[35] = 2.65, p < 0.05, d = 0.44$, 95% CI [0.48, 3.61] and CP2 ($t[35] = 3.12, p < .01, d = 0.52$, 95% CI [0.77, 3.67]. The second ANOVA comparing Uncued truth and Small lie condition revealed an interaction between response and site, but no main effect of response (Table 3C-B1). Planned focal analyses showed a decreased LPC for lying compared with uncued truth at Pz ($t[35] = 2.34, p < .05, d = 0.39$, 95% CI [0.15, 2.17]).

An additional ANOVA comparing the Cued truth and Uncued truth conditions showed a main effect of response and an interaction between response and site (Table 3C-B1). Planned focal analyses showed a smaller LPC for uncued truth compared to cued truth at Pz ($t[35] = 3.09, p < 0.01, d = 0.52$, 95% CI [0.76, 3.64]), P3 ($t[35] = 3.16, p < 0.01, d = 0.53$, 95% CI [0.71, 3.26], P4 ($t[35] = 2.57, p < 0.05, d = 0.43$, 95% CI [0.37, 3.13], CP1 ($t[35] = 2.07, p < 0.05, d = 0.35$, 95% CI [0.03, 2.73] and CP2 ($t[35] = 2.17, p < 0.05, d = 0.36$, 95% CI [0.10, 2.91].

To summarize, a smaller LPC for lying was found compared with cued truth and uncued truth. In addition, uncued truths showed a smaller LPC than the cued truths.

2.1.6 *Bereitschaftspotential* (BP, Response-locked)

The first ANOVA comparing Cued truth and Small lie conditions revealed an interaction between response and time window ($F[1,35] = 24.38$, $\eta^2 = 0.41$, $p < .001$). Further analyses indicated that between 600 and 400 ms prior to the response, the BP amplitude was smaller for lying compared with cued truth ($F[1,35] = 6.88$, $\eta^2 = 0.16$, $p < .05$). Similarly, the second ANOVA comparing the Uncued truth and Small lie conditions revealed a smaller BP for lying compared to uncued truth between 600 and 400 ms prior to the response ($F[1,35] = 5.30$, $\eta^2 = 0.13$, $p < .05$) (see Figure 3-5b).

An additional ANOVA comparing the Cued truth and Uncued truth conditions showed no interaction between response and time window ($F[1,35] = 1.86$, $\eta^2 = 0.05$, $p = .18$) and there was no difference in BP amplitude between these two response types between 600 and 400 ms prior to the response ($F[1,35] = 0.88$, $\eta^2 = 0.03$, $p = .35$).

In summary, a smaller BP for lying compared with both Cued truths and Uncued truths was found and there was no difference between Cued truths and Uncued truths.

A weighted average was carried out to combine Small lies and Big lies into a “Total lies” category. The comparison between “Truth” and “Total lie” is shown in Table 3C-B2, and the results are comparable to those in the Small lie condition.

2.2 *Lie detection*

Clip onset

For the regional analyses, the midline ANOVA showed an interaction between group and judgement during 500 – 700 ms ($F[1,17] = 8.95$, $\eta^2 = 0.35$, $p < 0.01$ for 500 – 600 ms; $F[1,17] = 10.21$, $\eta^2 = 0.38$, $p < 0.01$ for 600 – 700 ms). The mean amplitudes were averaged across stimuli for follow-up analyses. Planned focal analyses showed that “true” judgements were more positive than “lie” judgements for the “good” group (all $t[17] > 3.09$, $p < 0.01$), but not for “bad” group (all $t[17] < 1.40$, $p > .179$).

The four-way lateral ANOVA showed an interaction between group and judgement between 500 and 700 ms ($F[1,17] = 7.47$, $\eta^2 = 0.31$, $p < 0.05$ between 500 and 600 ms; $F[1,17] = 8.27$, $\eta^2 = 0.33$, $p = 0.01$ between 600 and 700 ms). There was a significant main effect of region between 100 and 900 ms (all $F[3,51] > 6.64$, $\eta^2 > 0.28$, $p < 0.01$). Follow-up three-way ANOVAs (group, stimulus and judgement) analyses were carried out for each region. The central region revealed an interaction between group and judgment between 300 and 700 ms (all $F[1,17] > 4.90$, $\eta^2 > 0.22$, $p < 0.05$) and the posterior region also revealed an interaction between group and judgment between 500 and 700 ms (all $F[1,17] > 6.65$, $\eta^2 > 0.28$, $p < 0.05$). In contrast, the frontal and temporal regions showed no such interaction. No other effects were found. The mean amplitudes were averaged across stimuli in central and posterior region for follow-up analyses. Analyses showed that that "true" judgments were more positive than "lie" judgments for the "good" group at both central (all $t[17] > 2.65$, $p < 0.05$) and posterior region (all $t[17] > 2.78$, $p < 0.05$), but not for "bad" group (all $t[17] < 1.34$, $p > .199$) (see Figure 3C-A).

Audio onset

For the regional analyses, the midline ANOVA revealed neither main effects nor interactions in any time windows from 0 to 900 ms. The four-way lateral ANOVA showed a main effect of region between 600 and 900 ms (all $F[3,51] > 6.97$, $\eta^2 > 0.29$, $p < 0.01$). No other main effect or interaction was found. Follow-up three-way ANOVAs (group, stimulus and judgement) carried out for each region showed neither a main effect nor interactions.

In summary, the amplitude of visual components P1 and N1 showed no differences between groups. The regional analyses showed an interaction between group and judgement between 500 and 700 ms at midline, central and posterior regions. This interaction was driven by the difference of mean amplitude between "truth" and "lie" judgement (collapsing stimuli) in the "good" group. Finally, no differences were found between the two groups using ERP time-locked to audio onset.

2.3 *Comparison between the lie detection and head movement tasks*

Clip onset

The mean amplitude for the P1 and N1 were analysed with repeated-measures ANOVAs using 2 factors: Task (lie detection, head movement) and Stimulus (truth, lie). The main effect of task was significant for the N1 ($F[1,35] = 16.56$, $\eta^2 = 0.32$, $p < 0.001$) with the head movement task showing a larger N1 than the lie detection task. The main effect of task was nearly significant for the P1 ($F[1,35] = 3.77$, $\eta^2 = 0.10$, $p = 0.06$) with the lie detection task showing a larger P1 than the head movement task. There was no main effect of stimulus and no interaction between task and stimuli.

For the regional analyses, the midline ANOVA showed a main effect of task between 100 and 900 ms (all $F[1,35] > 5.84$, $\eta^2 = 0.14$, $p < 0.05$) where the mean amplitude was more negative for the head movement task than the detection task. There was no main effect of stimulus and no interaction between task and stimulus.

The three-way lateral ANOVA (factors: task, region and stimulus) showed an interaction between task and region between 400 and 600 ms ($F[3,105] = 6.88$, $\eta^2 = 0.16$, $p < 0.01$ between 400 and 500 ms; $F[3,105] = 4.47$, $\eta^2 = 0.11$, $p = 0.05$ between 500 and 600 ms).The significant main effect of task persisted between 100 and 900 ms (all $F[1,35] > 4.63$, $\eta^2 = 0.12$, $p < 0.05$) where the mean amplitude was more negative for the head movement task than the detection task. There was a main effect of region in all time windows (all $F[3,105] > 11.18$, $\eta^2 = 0.24$, $p < 0.001$). Follow-up two-way ANOVAs (task and stimulus) were carried out for each region. The first difference was a main effect of task between 100 and 300 ms in the posterior region (all $F[1,35] > 8.64$, $\eta^2 = 0.20$, $p < 0.01$) and reflected the modulation of visual components (i.e., P1 and N1) with the head movement task showing a more negative mean amplitude than the lie detection task. No main effect of task was found in later time windows in the posterior region. For central and temporal regions, the main effect of task persisted between 100 and 900 ms (all $F[1,35] > 4.73$, $\eta^2 = 0.12$, $p < 0.05$, except for the central region between 700 and 800 ms showing a nearly significant difference, $F[1,35] = 3.94$, $\eta^2 = 0.10$, $p = 0.055$) with the head movement task consistently showing more negative mean amplitude than the lie detection task. The frontal region only showed a main effect of task between 300 and 700 ms (all $F[1,35] > 5.40$, $\eta^2 = 0.13$, $p < 0.05$) with more negative amplitude for the head movement task. From the pattern of lateral ANOVA result, it appeared that the difference of mean amplitude between two tasks shifted from posterior to anterior region through the epoch time (see Figure 3-6c&d).

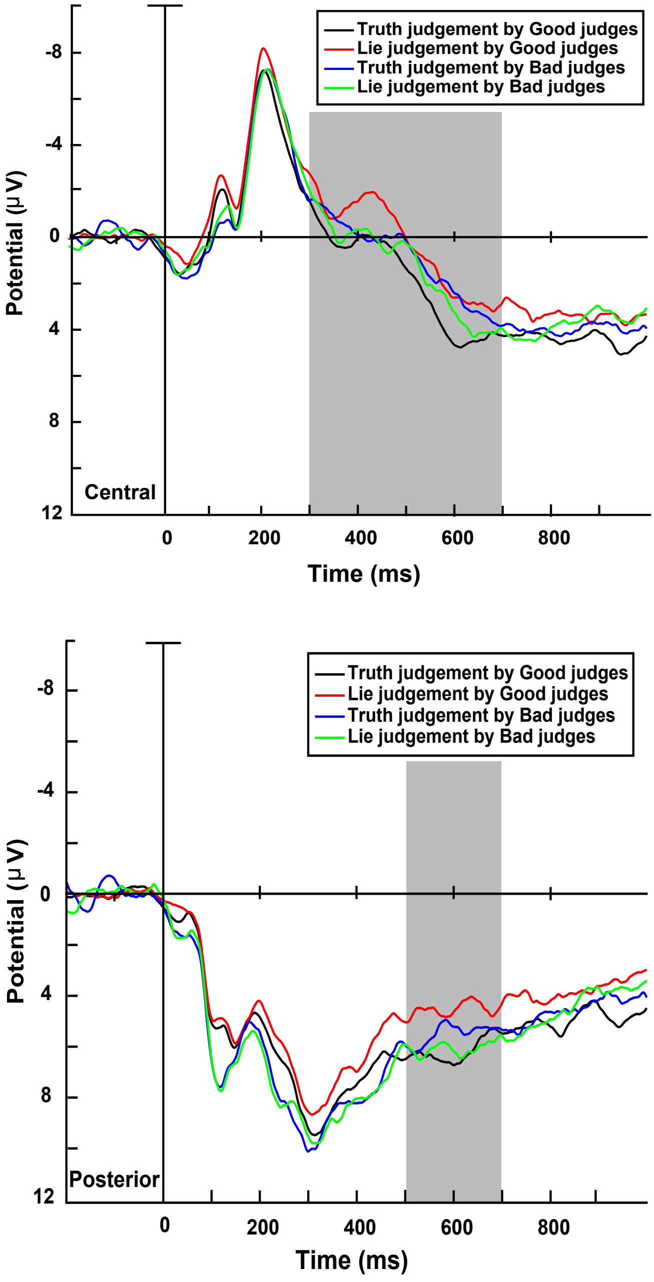
Audio onset

For the regional analyses, the midline ANOVA showed a main effect of task between 500 and 700 ms ($F[1,35] = 4.42$, $\eta^2 = 0.11$, $p < 0.05$ between 500 and 600 ms; $F[1,35] = 4.32$, $\eta^2 = 0.11$, $p < 0.05$ between 600 and 700 ms) where the mean amplitude was more negative for the lie detection than the head movement task. There was no main effect of stimulus and no interaction between task and stimulus.

The three-way lateral ANOVA showed an interaction between task and region between 200 and 600 ms (all $F[3,105] > 4.42$, $\eta^2 = 0.11$, $p < 0.05$). A significant main effect of region persisted between 200 and 900 ms (all $F[3,105] > 4.35$, $\eta^2 = 0.11$, $p < 0.05$). Follow-up two-way ANOVAs (task and stimulus) were carried out for each region. The significant main effect of task was only found in the posterior region between 400 and 700 ms (all $F[1,35] > 4.27$, $\eta^2 = 0.11$, $p < 0.05$) where the mean amplitude was more negative for the lie detection than the head movement task (see Figure 3C-B). No other significant effects were found in any other regional analyses.

To summarise, the head movement task showed a larger N1 than the lie detection task. The regional analyses suggested that the difference of amplitudes shifted from posterior to anterior region along the epoch time. In addition, no main effect of stimuli was found for both clip and audio onset.

Figure 3C-A. ERPs at (a) central and (b) posterior regions for trials judged as truthful and deceptive for both Good-judge and Bad-judge groups



(a)

(b)

Figure 3C-B. Topographic map showing the difference between the Lie detection and Head movement tasks between 400 and 700 ms (audio onset)

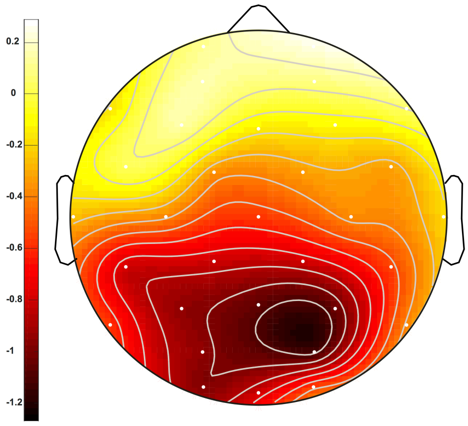


Table 3C-A1

Lateral (Lat) and Midline (Mid) ANOVAs on N200 (250, 350 ms) and P300 (450, 650 ms) amplitude comparing Cued Truth and Small Lie, Uncued Truth and Small Lie, and Cued Truth and Uncued Truth.

ERP	N200						P300					
	Lat			Mid			Lat			Mid		
Source	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2
Cued Truth vs Small Lie												
Response	0.08	.78	.00	0.35	.56	.01	**4.55**	**<.05**	**.12**	**9.37**	**<.01**	**.21**
R x Site	2.24	.07	.06	0.97	.38	.03	**5.54**	**=.001**	**.14**	**7.87**	**<.001**	**.18**
R x Hemi	**5.80**	**.02**	**.14**				0.03	.86	.00			
R x S x H	1.90	.08	.05				1.93	.07	.05			
Uncued Truth vs Small Lie												
Response	1.72	.20	.05	0.62	.44	.02	0.20	.66	.01	0.01	.92	.00
R x S	1.75	.15	.05	0.60	.57	.02	0.84	.47	.02	1.42	.25	.04
R x H	1.15	.29	.03				0.32	.57	.01			
R x S x H	0.94	.48	.03				0.90	.51	.03			
Cued Truth vs Uncued Truth												
Response	1.29	.26	.04	2.55	.12	.07	**8.69**	**<.01**	**.20**	**15.98**	**<.001**	**.31**
R x S	**2.90**	**.03**	**.08**	1.40	.25	.04	**6.73**	**<.001**	**.16**	**7.94**	**<.001**	**.19**
R x H	0.90	.35	.03				0.67	.42	.02			
R x S x H	1.72	.10	.05				1.83	.08	.05			

Note. Degrees of freedom: Response, Response x Hemisphere: 1, 35; Response x Site (Mid): 3, 105; Response x Site (Lat), Response x Hemisphere x Site: 13, 455. F values are with Greenhouse-Geisser correction.

Table 3C-A2

Lateral (Lat) and Midline (Mid) ANOVAs on N200 (250, 350 ms) and P300 (450, 650 ms) amplitude comparing Cued Truth and Total Lie, and Uncued Truth and Total Lie.

ERP	N200						P300					
	Lat			Mid			Lat			Mid		
Source	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2
Cued Truth vs Total Lie												
Response	0.33	.57	.01	0.06	.82	.00	3.26	.08	.09	**7.64**	**<.01**	**.18**
R x Site	1.54	.20	.04	0.47	.62	.01	**4.60**	**<.01**	**.12**	**6.98**	**<.01**	**.17**
R x Hemi	3.67	.06	.10				0.03	.87	.00			
R x S x H	1.86	.08	.05				1.99	.06	.05			
Uncued Truth vs Total Lie												
Response	2.79	.10	.07	1.78	.19	.05	0.67	.42	.02	0.26	.61	.01
R x S	1.05	.37	.03	0.39	.69	.01	0.72	.53	.02	1.08	.35	.03
R x H	0.39	.54	.01				0.39	.54	.01			
R x S x H	1.23	.29	.03				1.16	.33	.03			

Note. Degrees of freedom: Response, Response x Hemisphere: 1, 35; Response x Site (Mid): 3, 105; Response x Site (Lat), Response x Hemisphere x Site: 13, 455. F values are with Greenhouse-Geisser correction.

Table 3C-B1

Lateral (Lat) and Midline (Mid) ANOVAs on MFN (0, 100 ms), PRP (-200, 0 ms) and LPC (-250, 0 ms) amplitude comparing between Cued Truth and Small Lie, Uncued Truth and Small Lie, and Cued Truth and Uncued Truth. The site factor includes Fz, F3, F4, FC1, FC2, Cz for MFN and PRP; Pz, P3, P4, CP1, CP2 for LPC.

ERP	MFN			PRP			LPC		
Source	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2
				Cued Truth vs Small Lie					
Response	3.51	.07	.09	4.02	.05	.10	**11.58**	**<.01**	**.25**
R x Site	**3.02**	**<.05**	**.08**	**5.71**	**=.001**	**.14**	**5.67**	**=.001**	**.14**
				Uncued Truth vs Small Lie					
Response	0.06	.81	.00	0.00	.96	.00	2.13	.15	.06
R x Site	1.77	.14	.05	**3.06**	**<.05**	**.08**	**2.76**	**<.05**	**.07**
				Cued Truth vs Uncued Truth					
Response	2.80	.10	.07	4.09	.05	.11	**7.32**	**=.01**	**.17**
R x Site	2.34	.07	.06	**3.84**	**<.05**	**.10**	**2.80**	**<.05**	**.07**

Note. Degrees of freedom: Response: 1, 35; Response x Site (MFN&PRP): 5, 175; Response x Site (LPC): 4, 140. F values are with Greenhouse-Geisser correction.

Table 3C-B2

Lateral (Lat) and Midline (Mid) ANOVAs on MFN (0, 100 ms), PRP (-200, 0 ms) and LPC (-250, 0 ms) amplitude comparing Cued Truth and Total Lie, and Uncued Truth and Total Lie. The site factor includes Fz, F3, F4, FC1, FC2, Cz for MFN and PRP; Pz, P3, P4, CP1, CP2 for LPC.

ERP	MFN			PRP			LPC		
Source	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2
				Cued Truth vs Total Lie					
Response	2.02	.16	.06	3.17	.08	.08	**9.78**	**<.01**	**.22**
R x Site	**3.09**	**<.05**	**.08**	**6.23**	**=.001**	**.15**	**6.04**	**=.001**	**.15**
				Uncued Truth vs Total Lie					
Response	1.18	.28	.03	0.21	.65	.01	1.46	.24	.04
R x Site	2.05	.11	.06	**3.94**	**<.01**	**.10**	**3.13**	**<.05**	**.08**

Note. Degrees of freedom: Response: 1, 35; Response x Site (MFN&PRP): 5, 175; Response x Site (LPC): 4, 140. F values are with Greenhouse-Geisser correction.

APPENDIX 5A

Subject information and consent form for Study 4

DICHIARAZIONE DI CONSENSO INFORMATO

UNIVERSITÀ DEGLI STUDI DI PADOVA
DIPARTIMENTO DI PSICOLOGIA GENERALE

CORRELATI NEURALI DEL DECISION-MAIKING

Il/La sottoscritto/a ____________________________ dichiara:

- di essere stato/a messo/a a conoscenza delle procedure sperimentali relative all'indagine scientifica alla quale liberamente partecipa come volontario, al fine di contribuire all'avanzamento delle conoscenze nel campo delle funzioni cerebrali;

- di essere stato informato circa la possibilità di ritirarsi dalla ricerca in qualsiasi momento, senza motivazione, senza incorrere in alcuna penalizzazione ed ottenendo il non utilizzo dei suoi dati;

- di essere stato/a informato/a riguardo alle finalità e agli obiettivi della ricerca in questione;

- di aver preso visione diretta dell'ambiente in cui avverranno i rilievi sperimentali e degli apparati che saranno utilizzati a tale scopo;

- di essere stato messo a conoscenza che i risultati di tale ricerca, mantenendo l'anonimato dei soggetti partecipanti, potranno eventualmente essere comunicati ad altri ricercatori in occasione di congressi o riunioni scientifiche;

- di aver ricevuto soddisfacenti informazioni relativamente al principio di mantenimento della riservatezza delle informazioni relative e/o scaturite dall'esame della propria persona;

- di essere stato informato che le immagini verranno acquisite senza alcuna finalità diagnostica e di poter richiedere copia delle immagini in formato grezzo, per la cui interpretazione clinica potrò rivolgermi ad uno specialista di mia fiducia.

Si informa che tutti i dati personali a Lei relativi verranno trattati in conformità al Decreto Legislativo 30 giugno 2003 n. 196 “Codice in materia di protezione dei dati personali”. Si informa inoltre che tutti i risultati ottenuti dalle analisi connesse alle attività di ricerca o sperimentazione, così come ogni altro atto medico, sono da considerarsi strettamente confidenziali e sottoposti al vincolo del segreto professionale e della legislazione vigente in materia.

Padova, lì ________________________

Firma ____________________________

Firma dello sperimentatore che ha raccolto consenso _______________________.

Nel caso in cui il personale medico rilevasse informazioni di potenziale interesse per il mio stato di salute

☐ DESIDERO ☐ NON DESIDERO

essere contattato. In caso affermativo chiedo di essere contattato ai seguenti recapiti

TELEFONO............................ EMAIL..................................

LUOGO...................................... DATA..

NOME/COGNOME (STAMPATELLO)...

FIRMA..

LIST OF ABBREVIATIONS

ACC	Anterior Cingulate Cortex
AI	Agreement Index
ALE	Activation Likelihood Estimate
AnG	Angular Gyrus
ANS	Autonomic Nervous System
ATTA	Abbreviate Torrance Test for Adults
AUC	Area Under the Curve
BA	Brodmann's Area
BOLD	Blood-Oxygen-Level-Dependent
BP	Bereitschaftspotential
CAQ	Creativity Achievement Questionnaire
CIT	Concealed Information Test
CK	Concealed Knowledge
CM	Countermeasure
CNV	Contingent Negative-going Variation
CRN	Correct Response Negativity
DeceIT	Deceptive Interaction Task
DLPFC	Dorsolateral Prefrontal Cortex
EEG	Electroencephalogram
EMS	Episodic Memory Survey
EPD	Emotional Photos Description
ER	Error Rate
ERP	Event-Related Potential
fMRI	functional Magnetic Resonance Imaging
FWHM	Full-width at Half-maximum
GC	Middle Cingulate Gyrus
GFd	Medial/superior Frontal Gyrus
GKT	Guilty Knowledge Test
GLM	General Linear Models
HRF	Hemodynamic Response Function
ICA	Independent Component Analysis
IPL	Inferior Parietal Lobule
IRI	Interpersonal Reactivity Index
ITI	Inter-Trial Interval

LGFi	Left Inferior Frontal Gyrus
LGsm	Left Supramarginal Gyrus
LLPi	Left Inferior Parietal Lobule
LPC	Later Positive Component
MCSD	Marlowe Crowne Social Desirability Scale
MFN	Medial Frontal Negativity
MNI	Montreal Neurological Institute
NK	No Knowledge
OP	Opponent
OR	Orienting Reflex
OS	Opinion Survey
PHG	Parahippocampal Gyrus
PRP	Pre-response Positivity
RAT	Remote Associate Test
RGFi	Right Inferior Frontal Gyrus
rIFG	right Inferior Frontal Gyrus
RLPi	Right Inferior Parietal Lobule
ROC	Receiver Operating Characteristics
ROI	Region of Interest
rSMG	right Supramarginal Gyrus
RT	Response time
SCR	Skin Conductance Response
SDT	Signal Detection Theory
STS	Superior Temporal Sulcus
SVM	Support Vector Machine
TAS	Toronto Alexithymia Scale
Thal	Thalamus
ToM	Theory of Mind
TP	Temporal Pole
TPJ	Temporo-Parietal Junction
VLPFC	Ventrolateral Prefrontal Cortex
WASI	Wechsler Abbreviated Scale of Intelligence

Reference

Abe, N., Okuda, J., Suzuki, M., Sasaki, H., Matsuda, T., Mori, E., . . . Fujii, T. (2008). Neural correlates of true memory, false memory, and deception. *Cerebral Cortex, 18*(12), 2811-2819.

Abe, N., Suzuki, M., Mori, E., Itoh, M., & Fujii, T. (2007). Deceiving others: distinct neural responses of the prefrontal cortex and amygdala in simple fabrication and deception with social interactions. *Journal of cognitive neuroscience, 19*(2), 287-295.

Abe, N., Suzuki, M., Tsukiura, T., Mori, E., Yamaguchi, K., Itoh, M., & Fujii, T. (2006). Dissociable roles of prefrontal and anterior cingulate cortices in deception. *Cerebral Cortex, 16*(2), 192-199.

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends Cogn Sci, 4*(7), 267-278.

Amabile, T. (1996). *Creativity in context*: Westview press.

Anderson, D. E., Ansfield, M. E., & DePaulo, B. M. (1999). Love's best habit: Deception in the context of relationships.

Arden, R., Chavez, R. S., Grazioplene, R., & Jung, R. E. (2010). Neuroimaging creativity: a psychometric view. *Behavioural brain research, 214*(2), 143-156.

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends Cogn Sci, 8*(4), 170-177.

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2014). Inhibition and the right inferior frontal cortex: one decade on. *Trends Cogn Sci, 18*(4), 177-185.

Baumgartner, T., Fischbacher, U., Feierabend, A., Lutz, K., & Fehr, E. (2009). The neural circuitry of a broken promise. *Neuron, 64*(5), 756-770.

Beaussart, M. L., Andrews, C. J., & Kaufman, J. C. (2013). Creative liars: The relationship between creativity and integrity. *Thinking Skills and Creativity, 9*, 129-134.

Ben-Shakhar, G. (2011). 11 Countermeasures. *Memory detection: Theory and application of the concealed information test*, 200.

Ben-Shakhar, G., & Elaad, E. (2003). The validity of psychophysiological detection of information with the Guilty Knowledge Test: a meta-analytic review: American Psychological Association.

Ben-Shakhar, G., & Furedy, J. J. (2012). *Theories and applications in the detection of deception: A psychophysiological and international perspective*: Springer Science & Business Media.

Benedek, M., Könen, T., & Neubauer, A. C. (2012). Associative abilities underlying creativity. *Psychology of Aesthetics, Creativity, and the Arts, 6*(3), 273.

Berkowitz, A. L., & Ansari, D. (2010). Expertise-related deactivation of the right temporoparietal junction during musical improvisation. *Neuroimage, 49*(1), 712-719.

Berlad, I., & Pratt, H. (1995). P300 in response to the subject's own name. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section, 96*(5), 472-474.

Berlyne, D. E. (1960). Conflict, arousal, and curiosity.

Bhatt, S., Mbwana, J., Adeyemo, A., Sawyer, A., Hailu, A., & Vanmeter, J. (2009). Lying about facial recognition: an fMRI study. *Brain and cognition, 69*(2), 382-390.

Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and social psychology Review, 10*(3), 214-234.

Bond Jr, C. F., & DePaulo, B. M. (2008). Individual differences in judging deception: accuracy and bias. *Psychol Bull, 134*(4), 477.

Bond Jr, C. F., & Robinson, M. (1988). The evolution of deception. *Journal of Nonverbal Behavior, 12*(4), 295-307.

Botvinick, M., Nystrom, L. E., Fissell, K., Carter, C. S., & Cohen, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature, 402*(6758), 179.

Bowden, E. M., & Jung-Beeman, M. (2003). Normative data for 144 compound remote associate problems. *Behav Res Methods Instrum Comput, 35*(4), 634-639.

Bowman, H., Filetti, M., Alsufyani, A., Janssen, D., & Su, L. (2014). Countering countermeasures: detecting identity lies by detecting conscious breakthrough. *PloS one, 9*(3), e90595.

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. *Cereb Cortex, 11*(9), 825-836.

Brenkert, G. G. (2009). Innovation, rule breaking and the ethics of entrepreneurship. *Journal of Business Venturing, 24*(5), 448-464.

Brooks, J. L., Zoumpoulaki, A., & Bowman, H. (2017). Data-driven region-of-interest selection without inflating Type I error rate. *Psychophysiology, 54*(1), 100-113.

Brophy, D. R. (1998). Understanding, measuring, and enhancing individual creative problem-solving efforts. *Creativity Research Journal, 11*(2), 123-150.

Brophy, D. R. (2001). Comparing the attributes, activities, and performance of divergent, convergent, and combination thinkers. *Creativity Research Journal, 13*(3-4), 439-455.

Brunia, C. H., van Boxtel, G. J., & Böcker, K. B. (2012). Negative slow waves as indices of anticipation: the Bereitschaftspotential, the contingent negative variation, and the stimulus-preceding negativity.

Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication theory, 6*(3), 203-242.

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends Cogn Sci, 4*(6), 215-222.

Carlson, S. M., Moses, L. J., & Hix, H. R. (1998). The role of inhibitory processes in young children's difficulties with deception and false belief. *Child development, 69*(3), 672-691.

Carson, S. H., Peterson, J. B., & Higgins, D. M. (2005). Reliability, Validity, and Factor Structure of the Creative Achievement Questionnaire. *Creativity Research Journal, 17*(1), 37-50. doi:10.1207/s15326934crj1701_4

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science, 280*(5364), 747-749.

Chang, C. C., & Lin, C. J. (2011). LIBSVM: a library for support vector machines. *ACM Trans. Intell. Syst. Technol., 2*, 1-27.

Chang, L. J., Yarkoni, T., Khaw, M. W., & Sanfey, A. G. (2013). Decoding the role of the insula in human cognition: functional parcellation and large-scale reverse inference. *Cereb Cortex, 23*(3), 739-749. doi:10.1093/cercor/bhs065

Chen, A., Xu, P., Wang, Q., Luo, Y., Yuan, J., Yao, D., & Li, H. (2008). The timing of cognitive control in partially incongruent categorization. *Human brain mapping, 29*(9), 1028-1039.

Christ, S. E., Van Essen, D. C., Watson, J. M., Brubaker, L. E., & McDermott, K. B. (2009). The contributions of prefrontal cortex and executive control to deception: evidence from activation likelihood estimate meta-analyses. *Cerebral Cortex, 19*(7), 1557-1566.

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The reorienting system of the human brain: from environment to theory of mind. *Neuron, 58*(3), 306-324.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci, 3*(3), 201-215. doi:10.1038/nrn755

Cropley, D. H., Kaufman, J. C., & Cropley, A. J. (2008). Malevolent creativity: A functional model of creativity in terrorism and crime. *Creativity Research*

Journal, 20(2), 105-115.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *J Consult Psychol, 24*, 349-354.

Cui, Q., Vanman, E. J., Wei, D., Yang, W., Jia, L., & Zhang, Q. (2014). Detection of deception based on fMRI activation patterns underlying the production of a deceptive response and receiving feedback about the success of the deception after a mock murder crime. *Soc Cogn Affect Neurosci, 9*(10), 1472-1480. doi:10.1093/scan/nst134

Dale, A. (1999). Optimal experimental design for event-related fMRI. *Hum Brain Mapp, 8*, 109-114.

Davatzikos, C., Ruparel, K., Fan, Y., Shen, D., Acharyya, M., Loughead, J., . . . Langleben, D. D. (2005). Classifying spatial patterns of brain activity with machine learning methods: application to lie detection. *Neuroimage, 28*(3), 663-668.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *J Pers Soc Psychol, 44*(1), 113.

Dawson, M. E., Filion, D. L., & Schell, A. M. (1989). Is elicitation of the autonomic orienting response associated with allocation of processing resources? *Psychophysiology, 26*(5), 560-572.

DePaulo, B. M., & Kashy, D. A. (1998). Everyday lies in close and casual relationships. *J Pers Soc Psychol, 74*(1), 63.

DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *J Pers Soc Psychol, 70*(5), 979.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychol Bull, 129*(1), 74.

DePaulo, B. M., Zuckerman, M., & Rosenthal, R. (1980). Humans as lie detectors. *Journal of Communication, 30*(2), 129.

Derrfuss, J., Brass, M., Neumann, J., & Von Cramon, D. Y. (2005). Involvement of the inferior frontal junction in cognitive control: Meta-analyses of switching and Stroop studies. *Human brain mapping, 25*(1), 22-34.

Donchin, E., & Coles, M. G. (1988). Is the P300 component a manifestation of context updating. *Behavioral and brain sciences, 11*(3), 357-427.

Dong, G., Wu, H., & Lu, Q. (2010). Attempting to hide our real thoughts: Electrophysiological evidence from truthful and deceptive responses during evaluation. *Neuroscience letters, 479*(1), 1-5.

Ekman, P. (2009). *Telling Lies: Clues to Deceit in the Marketplace, Politics, and*

Marriage (Revised Edition): WW Norton & Company.

Ekman, P., & O'sullivan, M. (1991). Who can catch a liar? *American psychologist, 46*(9), 913.

Elaad, E., & Ben-Shakhar, G. (1991). Effects of mental countermeasures on psychophysiological detection in the guilty knowledge test. *International Journal of Psychophysiology, 11*(2), 99-108.

Elton, M., Schandry, R., & Sparrer, B. (1983). A comparative investigation of ERP components and the SCR in a habituation and dishabituation paradigm. *International Journal of Neuroscience, 22*(1-2), 55-62.

Eysenck, H. J. (1995). *Genius: The natural history of creativity* (Vol. 12): Cambridge University Press.

Falkenstein, M., Hohnsbein, J., & Hoormann, J. (1994). Effects of choice complexity on different subcomponents of the late positive complex of the event-related potential. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section, 92*(2), 148-160.

Falkenstein, M., Hohnsbein, J., & Hoormann, J. (1995). Event-related potential correlates of errors in reaction tasks. *Electroencephalography and clinical neurophysiology. Supplement, 44*, 287-296.

Falkenstein, M., Hohnsbein, J., Hoormann, J., & Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalography and clinical neurophysiology, 78*(6), 447-455.

Fang, F., Liu, Y., & Shen, Z. (2003). Lie detection with contingent negative variation. *International Journal of Psychophysiology, 50*(3), 247-255.

Farwell, L. A., & Donchin, E. (1991). The Truth Will Out: Interrogative Polygraphy ("Lie Detection") With Event-Related Brain Potentials. *Psychophysiology, 28*(5), 531-547.

Feist, G. J. (1998). A meta-analysis of personality in scientific and artistic creativity. *Personality and social psychology Review, 2*(4), 290-309.

Feist, G. J. (1999). 14 The Influence of Personality on Artistic and Scientific Creativity. *Handbook of creativity*, 273.

Filion, D. L., Dawson, M. E., Schell, A. M., & Hazlett, E. A. (1991). The relationship between skin conductance orienting the allocation of processing resources. *Psychophysiology, 28*(4), 410-424.

Finke, R. A., Ward, T. B., & Smith, S. M. (1992). Creative cognition: Theory, research, and applications.

Fletcher, P. C., Palomero-Gallagher, N., Zafiris, O., Fink, G. R., Tyler, L. K., & Zilles, K. (2002). The influence of explicit instructions and stimulus material on lateral frontal responses to an encoding task. *Neuroimage, 17*(2), 780-791.

Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology, 45*(1), 152-170.

Folstein, J. R., Van Petten, C., & Rose, S. A. (2008). Novelty and conflict in the categorization of complex stimuli. *Psychophysiology, 45*(3), 467-479.

Frith, C. (2000). The role of dorsolateral prefrontal cortex in the selection of action as revealed by functional imaging. *Control of cognitive processes: Attention and performance XVIII*, 549-565.

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*(12), 1287-1301.

Gamer, M., & Berti, S. (2010). Task relevance and recognition of concealed information have different influences on electrodermal activity and event-related brain potentials. *Psychophysiology, 47*(2), 355-364.

Gamer, M., & Berti, S. (2012). P300 amplitudes in the concealed information test are less affected by depth of processing than electrodermal responses. *Frontiers in human neuroscience, 6*.

Gamer, M., Klimecki, O., Bauermann, T., Stoeter, P., & Vossel, G. (2009). fMRI-activation patterns in the detection of concealed information rely on memory-related effects. *Social cognitive and affective neuroscience, 7*(5), 506-515.

Gamer, M., Klimecki, O., Bauermann, T., Stoeter, P., & Vossel, G. (2012). fMRI-activation patterns in the detection of concealed information rely on memory-related effects. *Soc Cogn Affect Neurosci, 7*(5), 506-515. doi:10.1093/scan/nsp005

Ganis, G., Bridges, D., Hsu, C.-W., & Schendan, H. E. (2016). Is anterior N2 enhancement a reliable electrophysiological index of concealed information? *Neuroimage, 143*, 152-165.

Ganis, G., & Keenan, J. P. (2009). The cognitive neuroscience of deception. *Soc Neurosci, 4*(6), 465-472. doi:10.1080/17470910802507660

Ganis, G., Kosslyn, S. M., Stose, S., Thompson, W., & Yurgelun-Todd, D. A. (2003). Neural correlates of different types of deception: an fMRI investigation. *Cerebral Cortex, 13*(8), 830-836.

Ganis, G., Rosenfeld, J. P., Meixner, J., Kievit, R. A., & Schendan, H. E. (2011). Lying in the scanner: covert countermeasures disrupt deception detection by functional magnetic resonance imaging. *Neuroimage, 55*(1), 312-319. doi:10.1016/j.neuroimage.2010.11.025

Ganis, G., & Schendan, H. E. (2012). Concealed semantic and episodic autobiographical memory electrified. *Frontiers in human neuroscience, 6*.

García-Larrea, L., & Cézanne-Bert, G. (1998). P3, positive slow wave and working memory load: a study on the functional correlates of slow wave activity. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section, 108*(3), 260-273.

Gati, I., & Ben-Shakhar, G. (1990). Novelty and significance in orientation and habituation: A feature-matching approach. *Journal of Experimental Psychology: General, 119*(3), 251.

Gehring, W. J., Goss, B., Coles, M. G., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological science, 4*(6), 385-390.

Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science, 295*(5563), 2279-2282.

Geng, J. J., & Vossel, S. (2013). Re-evaluating the role of TPJ in attentional control: contextual updating? *Neurosci Biobehav Rev, 37*(10 Pt 2), 2608-2620. doi:10.1016/j.neubiorev.2013.08.010

Gino, F., & Ariely, D. (2012). The dark side of creativity: original thinkers can be more dishonest. *J Pers Soc Psychol, 102*(3), 445.

Gino, F., & Wiltermuth, S. S. (2014). Evil genius? How dishonesty can lead to greater creativity. *Psychological science*, 0956797614520714.

Goff, K. (2002). *Abbreviated torrance test for adults*: Scholastic Testing Service Bensenville, IL.

Gombos, V. A. (2006). The cognition of deception: the role of executive processes in producing lies. *Genetic, social, and general psychology monographs, 132*(3), 197-214.

Good, P. L. (2005). *Resampling methods: 3d Edition*: Birkhauser.

Gratton, G., Coles, M. G., & Donchin, E. (1992). Optimizing the use of information: strategic control of activation of responses. *Journal of Experimental Psychology: General, 121*(4), 480.

Green, D., & Swets, J. (1966). Signal detection theory and psychophysics. 1966. *New York, 888*, 889.

Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron, 44*(2), 389-400.

Greene, J. D., & Paxton, J. M. (2009). Patterns of neural activity associated with honest and dishonest moral decisions. *Proceedings of the National Academy of Sciences, 106*(30), 12506-12511.

Grèzes, J., Frith, C., & Passingham, R. E. (2004). Brain mechanisms for inferring deceit in the actions of others. *Journal of Neuroscience, 24*(24), 5500-5505.

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011a). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology, 48*(12), 1711-1725.

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011b). Mass univariate analysis of event-related brain potentials/fields II: Simulation studies. *Psychophysiology, 48*(12), 1726-1737.

Guilford, J. P. (1950). Creativity. *American psychologist, 5*, 444.

Guilford, J. P. (1967). The nature of human intelligence.

Hala, S., & Russell, J. (2001). Executive control within strategic deception: a window on early cognitive development? *Journal of experimental child psychology, 80*(2), 112-141.

Hanley, J. A., & Hajian-Tilaki, K. O. (1997). Sampling variability of nonparametric estimates of the areas under receiver operating characteristic curves: an update. *Academic radiology, 4*(1), 49-58.

Hanley, J. A., & McNeil, B. J. (1983). A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology, 148*, 839-843.

Hillyard, S. A., & Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proceedings of the National Academy of Sciences, 95*(3), 781-787.

Hocking, J. E., & Leathers, D. G. (1980). Nonverbal indicators of deception: A new theoretical perspective. *Communications Monographs, 47*(2), 119-131.

Hommel, B., Todd, P., Hills, T., & Robbins, T. (2012). Convergent and divergent operations in cognitive search. *Cognitive search: Evolution, algorithms, and the brain*, 221-235.

Honts, C. R., Devitt, M. K., Winbush, M., & Kircher, J. C. (1996). Mental and physical countermeasures reduce the accuracy of the concealed knowledge test. *Psychophysiology, 33*(1), 84-92.

Hsu, C. W., Chang, C. C., & Lin, C. J. (2003). *A Practical Guide to Support Vector*

Classification. Technical Report. Department of Computer Science. National Taiwan University. Retrieved from http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf

Hu, X., Pornpattananangkul, N., & Nusslock, R. (2015). Executive control-and reward-related neural processes associated with the opportunity to engage in voluntary dishonest moral decision making. *Cognitive, Affective, & Behavioral Neuroscience, 15*(2), 475-491.

Hu, X., Pornpattananangkul, N., & Rosenfeld, J. P. (2013). N200 and P300 as orthogonal and integrable indicators of distinct awareness and recognition processes in memory detection. *Psychophysiology, 50*(5), 454-464.

Hu, X., Wu, H., & Fu, G. (2011). Temporal course of executive control when lying about self-and other-referential information: an ERP study. *Brain Research, 1369*, 149-157.

Iidaka, T., Sadato, N., Yamada, H., & Yonekura, Y. (2000). Functional asymmetry of human prefrontal cortex in verbal and non-verbal episodic memory as revealed by fMRI. *Brain Res Cogn Brain Res, 9*(1), 73-83.

Isreal, J. B., Chesney, G. L., Wickens, C. D., & Donchin, E. (1980). P300 and tracking difficulty: Evidence for multiple resources in dual-task performance. *Psychophysiology, 17*(3), 259-273.

Isreal, J. B., Wickens, C. D., Chesney, G. L., & Donchin, E. (1980). The event-related brain potential as an index of display-monitoring workload. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 22*(2), 211-224.

Ito, A., Abe, N., Fujii, T., Hayashi, A., Ueno, A., Mugikura, S., . . . Mori, E. (2012). The contribution of the dorsolateral prefrontal cortex to the preparation for deception and truth-telling. *Brain Research, 1464*, 43-52.

Ito, A., Abe, N., Fujii, T., Ueno, A., Koseki, Y., Hashimoto, R., . . . Mori, E. (2011). The role of the dorsolateral prefrontal cortex in deception when remembering neutral and emotional events. *Neuroscience research, 69*(2), 121-128.

Jin, B., Strasburger, A., Laken, S. J., Kozel, F. A., Johnson, K. A., George, M. S., & Lu, X. (2009). Feature selection for fMRI-based deception detection. *BMC Bioinformatics, 10 Suppl 9*, S15. doi:10.1186/1471-2105-10-S9-S15

Johnson, R. (1988). The amplitude of the P300 component of the event-related potential: Review and synthesis. *Advances in psychophysiology, 3*, 69-137.

Johnson, R. (1993). On the neural generators of the P300 component of the event-related potential. *Psychophysiology, 30*, 90-90.

Johnson, R. (2014). The neural basis of deception and credibility assessment: a cognitive neuroscience perspective. *Credibility Assessment: Scientific Research and Applications, Elsevier, Amsterdam*, 217-300.

Johnson, R., Barnhardt, J., & Zhu, J. (2003). The deceptive response: effects of response conflict and strategic monitoring on the late positive component and episodic memory-related brain activity. *Biological psychology, 64*(3), 217-253.

Johnson, R., Barnhardt, J., & Zhu, J. (2004). The contribution of executive processes to deceptive responding. *Neuropsychologia, 42*(7), 878-901.

Johnson, R., Barnhardt, J., & Zhu, J. (2005). Differential effects of practice on the executive processes used for truthful and deceptive responses: an event-related brain potential study. *Cognitive Brain Research, 24*(3), 386-404.

Johnson, R., Henkell, H., Simon, E., & Zhu, J. (2008). The self in conflict: the role of executive processes during truthful and deceptive responses about attitudes. *Neuroimage, 39*(1), 469-482.

Johnston, V. S., Miller, D. R., & Burleson, M. H. (1986). Multiple P3s to emotional stimuli and their theoretical significance. *Psychophysiology, 23*(6), 684-694.

Karim, A. A., Schneider, M., Lotze, M., Veit, R., Sauseng, P., Braun, C., & Birbaumer, N. (2010). The truth about lying: inhibition of the anterior prefrontal cortex improves deceptive behavior. *Cerebral Cortex, 20*(1), 205-213.

Kaufman, J. C., Kaufman, S. B., & Lichtenberger, E. O. (2011). Finding creative potential on intelligence tests via divergent production. *Canadian Journal of School Psychology, 26*(2), 83-106.

Kelly, A., Hester, R., Murphy, K., Javitt, D. C., Foxe, J. J., & Garavan, H. (2004). Prefrontal-subcortical dissociations underlying inhibitory control revealed by event-related fMRI. *European Journal of Neuroscience, 19*(11), 3105-3112.

Kim, K. H. (2008). Meta-analyses of the relationship of creative achievement to both IQ and divergent thinking test scores. *The Journal of Creative Behavior, 42*(2), 106-130.

Kireev, M., Pakhomov, S., & Medvedev, S. (2007). Study of brain mechanisms in error detection during deceptive responses in the normal state and under the influence of alcohol. *Fiziologiia cheloveka, 34*(2), 13-22.

Kireev, M., Pakhomov, S., & Medvedev, S. (2008). Cerebral mechanisms of error detection during deceptive responses in the normal state and under the influence of alcohol. *Human Physiology, 34*(2), 141-149.

Kireev, M., Starchenko, M., Pakhomov, S., & Medvedev, S. (2007). Stages of the cerebral mechanisms of deceptive responses. *Human Physiology, 33*(6), 659-666.

Koechlin, E., & Hyafil, A. (2007). Anterior prefrontal function and the limits of human decision-making. *Science, 318*(5850), 594-598.

Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science, 302*(5648), 1181-1185.

Koechlin, E., & Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. *Trends Cogn Sci, 11*(6), 229-235.

Kok, & De Jong, H. L. (1980). Components of the event-related potential following degraded and undegraded visual stimuli. *Biological psychology, 11*(2), 117-133.

Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology, 38*(3), 557-577.

Kozel, F. A., Johnson, K. A., Mu, Q., Grenesko, E. L., Laken, S. J., & George, M. S. (2005). Detecting deception using functional magnetic resonance imaging. *Biological psychiatry, 58*(8), 605-613.

Kramer, A. F., Wickens, C. D., & Donchin, E. (1985). Processing of stimulus properties: Evidence for dual-task integrality. *Journal of Experimental Psychology: Human Perception and Performance, 11*(4), 393.

Kraut, R. E. (1978). Verbal and nonverbal cues in the perception of lying. *J Pers Soc Psychol, 36*(4), 380.

Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S., & Baker, C. I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nature Neuroscience, 12*(5), 535-540.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual*: NIMH, Center for the Study of Emotion & Attention.

Langleben, D. D., Hakun, J. G., Seelig, D., Wang, A.-L., Ruparel, K., Bilker, W. B., & Gur, R. C. (2016). Polygraphy and Functional Magnetic Resonance Imaging in Lie Detection: A Controlled Blind Comparison Using the Concealed Information Test. *The Journal of clinical psychiatry, 77*(10), 1372-1380.

Langleben, D. D., Schroeder, L., Maldjian, J., Gur, R., McDonald, S., Ragland, J. D., . . . Childress, A. R. (2002). Brain activity during simulated deception: an event-related functional magnetic resonance study. *Neuroimage, 15*(3), 727-732.

Lee, C. S., Huggins, A. C., & Therriault, D. J. (2014). A measure of creativity or intelligence? Examining internal and external structure validity evidence of the Remote Associates Test. *Psychology of Aesthetics, Creativity, and the Arts, 8*(4), 446.

Lee, C. S., & Therriault, D. J. (2013). The cognitive underpinnings of creative thought:

A latent variable analysis exploring the roles of intelligence and working memory in three creative thinking processes. *Intelligence, 41*(5), 306-320.

Lee, T. M., Au, R. K., Liu, H.-L., Ting, K., Huang, C.-M., & Chan, C. C. (2009). Are errors differentiable from deceptive responses when feigning memory impairment? An fMRI study. *Brain and cognition, 69*(2), 406-412.

Lev, B. (2004). Sharpening the intangibles edge. *Harvard business review, 6*, 109-116.

Levine, T. R., Park, H. S., & McCornack, S. A. (1999). Accuracy in detecting truths and lies: Documenting the “veracity effect”. *Communications Monographs, 66*(2), 125-144.

Levy, B. J., & Wagner, A. D. (2011). Cognitive control and right ventrolateral prefrontal cortex: reflexive reorienting, motor inhibition, and action updating. *Ann N Y Acad Sci, 1224*, 40-62. doi:10.1111/j.1749-6632.2011.05958.x

Lisofsky, N., Kazzer, P., Heekeren, H. R., & Prehn, K. (2014). Investigating socio-cognitive processes in deception: a quantitative meta-analysis of neuroimaging studies. *Neuropsychologia, 61*, 113-122.

Lu, J. G., Quoidbach, J., Gino, F., Chakroff, A., Maddux, W. W., & Galinsky, A. D. (2017). The dark side of going abroad: How broad foreign experiences increase immoral behavior. *J Pers Soc Psychol, 112*(1), 1.

Luck, S. J., & Yard, S. A. H. (1995). The role of attention in feature detection and conjunction discrimination: An electrophysiological analysis. *International Journal of Neuroscience, 80*(1-4), 281-297.

Ludwig, A. M. (1995). *The price of greatness: Resolving the creativity and madness controversy*: Guilford Press.

Lykken, D. T. (1959). The GSR in the detection of guilt. *Journal of Applied Psychology, 43*(6), 385.

Lykken, D. T. (1960). The validity of the guilty knowledge technique: The effects of faking. *Journal of Applied Psychology, 44*(4), 258.

Lykken, D. T. (1998). *A tremor in the blood: Uses and abuses of the lie detector*: Plenum Press.

Mameli, F., Mrakic-Sposta, S., Vergari, M., Fumagalli, M., Macis, M., Ferrucci, R., . . . Priori, A. (2010). Dorsolateral prefrontal cortex specifically processes general– but not personal–knowledge deception: multiple brain networks for lying. *Behavioural brain research, 211*(2), 164-168.

Mar, R. A. (2011). The neural bases of social cognition and story comprehension. *Annu Rev Psychol, 62*, 103-134.

Mathalon, D. H., Bennett, A., Askari, N., Gray, E. M., Rosenbloom, M. J., & Ford, J. M.

(2003). Response-monitoring dysfunction in aging and Alzheimer's disease: an event-related potential study. *Neurobiology of aging, 24*(5), 675-685.

Matsuda, I., Nittono, H., & Allen, J. J. (2012). The current and future status of the concealed information test for field use. *Frontiers in Psychology, 3*, 532. doi:10.3389/fpsyg.2012.00532

McCornack, S. A., Parks, M. R., & McLaughlin, M. L. (1986). Deception detection and relationship development: The other side of trust. *Communication yearbook, 9*, 377-389.

McDermott, K. B., Szpunar, K. K., & Christ, S. E. (2009). Laboratory-based and autobiographical retrieval tasks differ substantially in their neural substrates. *Neuropsychologia, 47*(11), 2290-2298.

Mednick, S. A. (1962). The associative basis of the creative process. *Psychol Rev, 69*, 220-232.

Mednick, S. A., & Mednick, M. T. (1967). *Examiner's manual, Remote Associates Test: College and adult forms 1 and 2*: Houghton Mifflin.

Meijer, E. H., Selle, N. K., Elber, L., & Ben-Shakhar, G. (2014). Memory detection with the Concealed Information Test: A meta analysis of skin conductance, respiration, heart rate, and P300 data. *Psychophysiology, 51*(9), 879-904.

Mognon, A., Jovicich, J., Bruzzone, L., & Buiatti, M. (2011). ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology, 48*(2), 229-240.

Molnar-Szakacs, I., Iacoboni, M., Koski, L., & Mazziotta, J. C. (2005). Functional segregation within pars opercularis of the inferior frontal gyrus: evidence from fMRI studies of imitation and action observation. *Cereb Cortex, 15*(7), 986-994. doi:10.1093/cercor/bhh199

Monteleone, G. T., Phan, K. L., Nusbaum, H. C., Fitzgerald, D., Irick, J.-S., Fienberg, S. E., & Cacioppo, J. T. (2009). Detection of deception using fMRI: better than chance, but well below perfection. *Social neuroscience, 4*(6), 528-538.

Mumford, M. D., & Gustafson, S. B. (1988). Creativity syndrome: Integration, application, and innovation. *Psychol Bull, 103*(1), 27.

National Research Council. (2003). *The Polygraph and Lie Detection. Committee to Review the Scientific Evidence on the Polygraph.* Washington, DC: The National Academic Press.

Naylor, H., Halliday, R., Callaway, E., Yano, L., & Walton, P. (1987). P3 as an index of visual information processing. *Electroencephalography and clinical neurophysiology. Supplement, 40*, 235.

Nielsen, B. D., Pickett, C. L., & Simonton, D. K. (2008). Conceptual versus experimental creativity: Which works best on convergent and divergent thinking tasks? *Psychology of Aesthetics, Creativity, and the Arts, 2*(3), 131.

Nobre, A. C., Rao, A., & Chelazzi, L. (2006). Selective attention to specific features within objects: behavioral and electrophysiological evidence. *Journal of cognitive neuroscience, 18*(4), 539-561.

Nose, I., Murai, J. i., & Taira, M. (2009). Disclosing concealed information on the basis of cortical activations. *Neuroimage, 44*(4), 1380-1386.

Nunez, J. M., Casey, B., Egner, T., Hare, T., & Hirsch, J. (2005). Intentional false responding shares neural substrates with response conflict and cognitive control. *Neuroimage, 25*(1), 267-277.

Oliveira, C. M., & Levine, T. R. (2008). Lie acceptability: A construct and measure. *Communication Research Reports, 25*(4), 282-288.

Olson, I. R., Plotzker, A., & Ezzyat, Y. (2007). The enigmatic temporal pole: a review of findings on social and emotional processing. *Brain, 130*(7), 1718-1731.

Panasiti, M. S., Pavone, E. F., Mancini, A., Merla, A., Grisoni, L., & Aglioti, S. M. (2014). The motor cost of telling lies: Electrocortical signatures and personality foundations of spontaneous deception. *Social neuroscience, 9*(6), 573-589.

Panasiti, M. S., Pavone, E. F., Merla, A., & Aglioti, S. M. (2011). Situational and dispositional determinants of intentional deceiving. *PloS one, 6*(4), e19465.

Parker, J. D., Taylor, G. J., & Bagby, R. M. (2001). The relationship between emotional intelligence and alexithymia. *Personality and Individual differences, 30*(1), 107-115.

Peskin, J. (1992). Ruse and representations: On children's ability to conceal information. *Developmental Psychology, 28*(1), 84.

Peth, J., Sommer, T., Hebart, M. N., Vossel, G., Buchel, C., & Gamer, M. (2015). Memory detection using fMRI - does the encoding context matter? *Neuroimage, 113*, 164-174. doi:10.1016/j.neuroimage.2015.03.051

Peth, J., Suchotzki, K., & Gamer, M. (2016). Influence of countermeasures on the validity of the Concealed Information Test. *Psychophysiology, 53*(9), 1429-1440.

Pfister, R., Foerster, A., & Kunde, W. (2014). Pants on fire: The electrophysiological signature of telling a lie. *Social neuroscience, 9*(6), 562-572.

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 Tesla1. *Academic radiology, 12*(2), 164-172.

Plucker, J. A. (1999). Is the proof in the pudding? Reanalyses of Torrance's (1958 to present) longitudinal data. *Creativity Research Journal, 12*(2), 103-114.

Polich, J. (1989). Habituation of P300 from auditory stimuli. *Psychobiology, 17*(1), 19-28.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical neurophysiology, 118*(10), 2128-2148.

Polich, J. (2012). Neuropsychology of P300. *Oxford handbook of event-related potential components*, 159-188.

Rissman, J., Greely, H. T., & Wagner, A. D. (2010). Detecting individual memories through the neural decoding of memory states and past experience. *Proceedings of the National Academy of Sciences of the United States of America, 107*(21), 9849-9854. doi:10.1073/pnas.1001028107

Rosenfeld, J. P. (2011). P300 in detecting concealed information. *Memory detection: Theory and application of the concealed information test*, 63-89.

Rosenfeld, J. P., Ben-Shakhar, G., & Ganis, G. (2012). Physiologically based methods of concealed memory detection. *Memory and law*.

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*(1-2), 157-161.

Rosenfeld, J. P., Soskins, M., Bosh, G., & Ryan, A. (2004). Simple, effective countermeasures to P300-based tests of detection of concealed information. *Psychophysiology, 41*(2), 205-219.

Roth, W. T., Ford, J. M., & Kopell, B. S. (1978). Long-latency evoked potentials and reaction time. *Psychophysiology, 15*(1), 17-23.

Ruchkin, D. S., Munson, R., & Sutton, S. (1982). P300 and slow wave in a message consisting of two events. *Psychophysiology, 19*(6), 629-642.

Runco, M. A. (2004). Personal creativity and culture *Creativity: When east meets west* (pp. 9-21).

Runco, M. A. (2010). Creativity has no dark side. *The dark side of creativity*, 15-32.

Rushby, J. A., & Barry, R. J. (2009). Single-trial event-related potentials to significant stimuli. *International Journal of Psychophysiology, 74*(2), 120-131.

Rushworth, M. F., Walton, M. E., Kennerley, S. W., & Bannerman, D. M. (2004). Action sets and decisions in the medial frontal cortex. *Trends Cogn Sci, 8*(9), 410-417. doi:10.1016/j.tics.2004.07.009

Sarlo, M., Lotto, L., Manfrinati, A., Rumiati, R., Gallicchio, G., & Palomba, D. (2012). Temporal dynamics of cognitive–emotional interplay in moral decision-making.

Journal of cognitive neuroscience, 24(4), 1018-1029.

Sasaki, M., Hira, S., & Matsuda, T. (2001). Effects of a mental countermeasure on the physiological detection of deception using the event-related brain potentials. *Shinrigaku kenkyu: The Japanese journal of psychology, 72*(4), 322-328.

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: the role of the temporo-parietal junction in “theory of mind”. *Neuroimage, 19*(4), 1835-1842.

Saxe, R., Moran, J. M., Scholz, J., & Gabrieli, J. (2006). Overlapping and non-overlapping brain regions for theory of mind and self reflection in individual subjects. *Social cognitive and affective neuroscience, 1*(3), 229-234.

Shibasaki, H., & Hallett, M. (2006). What is the Bereitschaftspotential? *Clinical neurophysiology, 117*(11), 2341-2356.

Shomstein, S. (2012). Cognitive functions of the posterior parietal cortex: top-down and bottom-up attentional control. *Front Integr Neurosci, 6*, 38. doi:10.3389/fnint.2012.00038

Shulman, G. L., McAvoy, M. P., Cowan, M. C., Astafiev, S. V., Tansy, A. P., d'Avossa, G., & Corbetta, M. (2003). Quantitative analysis of attention and detection signals during visual search. *Journal of neurophysiology, 90*(5), 3384-3397.

Siddle, D. A. (1991). Orienting, habituation, and resource allocation: An associative analysis. *Psychophysiology, 28*(3), 245-259.

Siddle, D. A., & Spinks, J. A. (1992). Orienting, habituation, and the allocation of processing resources. *Attention and information processing in infants and adults: Perspectives from human and animal research*, 227-262.

Silvia, P. J., Kaufman, J. C., & Pretz, J. E. (2009). Is creativity domain-specific? Latent class models of creative accomplishments and creative self-descriptions. *Psychology of Aesthetics, Creativity, and the Arts, 3*(3), 139.

Sip, K. E., Lynge, M., Wallentin, M., McGregor, W. B., Frith, C. D., & Roepstorff, A. (2010). The production and detection of deception in an interactive game. *Neuropsychologia, 48*(12), 3619-3626.

Sip, K. E., Roepstorff, A., McGregor, W., & Frith, C. D. (2008). Detecting deception: the scope and limits. *Trends Cogn Sci, 12*(2), 48-53.

Sip, K. E., Skewes, J. C., Marchant, J. L., McGregor, W. B., Roepstorff, A., & Frith, C. D. (2012). What if I get busted? Deception, choice, and decision-making in social interaction.

Sokolov, E. (1966). Orienting reflex as information regulator. *Psychological research in the USSR, 1*, 334-360.

Sokolov, E. N. (1963). Perception and the conditioned reflex.

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*(13), 2849-2853.

Spence, S. A., Kaylor-Hughes, C., Farrow, T. F., & Wilkinson, I. D. (2008). Speaking of secrets and lies: the contribution of ventrolateral prefrontal cortex to vocal deception. *Neuroimage, 40*(3), 1411-1418.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behav Res Methods Instrum Comput, 31*(1), 137-149.

Sternberg, R. J. (1988). 5 A three-facet model of creativity. *The nature of creativity: Contemporary psychological perspectives*, 125.

Sternberg, R. J. (1999). *Handbook of creativity*: Cambridge University Press.

Suchotzki, K., Crombez, G., Smulders, F. T., Meijer, E., & Verschuere, B. (2015). The cognitive mechanisms underlying deception: An event-related potential study. *International Journal of Psychophysiology, 95*(3), 395-405.

Suchotzki, K., Verschuere, B., Peth, J., Crombez, G., & Gamer, M. (2015). Manipulating item proportion and deception reveals crucial dissociation between behavioral, autonomic, and neural indices of concealed information. *Human brain mapping, 36*(2), 427-439.

Sun, S.-Y., Mai, X., Liu, C., Liu, J.-Y., & Luo, Y.-J. (2011). The processes leading to deception: ERP spatiotemporal principal component analysis and source analysis. *Social neuroscience, 6*(4), 348-359.

Sutton, S., Braren, M., Zubin, J., & John, E. (1965). Evoked-potential correlates of stimulus uncertainty. *Science, 150*(3700), 1187-1188.

Swick, D., Ashley, V., & Turken, A. U. (2008). Left inferior frontal gyrus is critical for response inhibition. *BMC Neurosci, 9*, 102. doi:10.1186/1471-2202-9-102

Taylor, G. J., Bagby, R. M., & Parker, J. D. (1992). The Revised Toronto Alexithymia Scale: some reliability, validity, and normative data. *Psychother Psychosom, 57*(1-2), 34-41.

Timm, H. W. (1982). Effect of altered outcome expectancies stemming from placebo and feedback treatments on the validity of the guilty knowledge technique. *Journal of Applied Psychology, 67*(4), 391.

Todd, J. J., Fougnie, D., & Marois, R. (2005). Visual short-term memory load suppresses temporo-parietal junction activity and induces inattentional blindness. *Psychological science, 16*(12), 965-972.

Tong, F., & Pratte, M. S. (2012). Decoding patterns of human brain activity. *Annu Rev*

Psychol, 63, 483-509. doi:10.1146/annurev-psych-120710-100412

Turken, U., & Swick, D. (1999). Response selection in the human anterior cingulate cortex. *Nature neuroscience, 2*(10), 920-924.

Ullsperger, M., & Von Cramon, D. Y. (2001). Subprocesses of performance monitoring: a dissociation of error processing and response competition revealed by event-related fMRI and ERPs. *Neuroimage, 14*(6), 1387-1401.

Uncapher, M. R., Boyd-Meredith, J. T., Chow, T. E., Rissman, J., & Wagner, A. D. (2015). Goal-Directed Modulation of Neural Memory Patterns: Implications for fMRI-Based Memory Detection. *J Neurosci, 35*(22), 8531-8545. doi:10.1523/JNEUROSCI.5145-14.2015

Van Veen, V., & Carter, C. S. (2002). The anterior cingulate as a conflict monitor: fMRI and ERP studies. *Physiology & behavior, 77*(4), 477-482.

Verschuere, B., Ben-Shakhar, G., & Meijer, E. (2011). *Memory detection: Theory and application of the Concealed Information Test*: Cambridge University Press.

Verschuere, B., Crombez, G., De Clercq, A., & Koster, E. H. (2004). Autonomic and behavioral responding to concealed information: Differentiating orienting and defensive responses. *Psychophysiology, 41*(3), 461-466.

Verschuere, B., & Shalvi, S. (2014). The truth comes naturally! Does it? *Journal of Language and Social Psychology, 33*(4), 417-423.

Verschueure, B., Ben-Shakhar, G., & Meijer, E. (Eds.). (2011). *Memory Detection*. Cambridge, UK: Cambridge University Press.

Vidal, F., Hasbroucq, T., Grapperon, J., & Bonnet, M. (2000). Is the 'error negativity'specific to errors? *Biological psychology, 51*(2), 109-128.

Vogel, E. K., & Luck, S. J. (2000). The visual N1 component as an index of a discrimination process. *Psychophysiology, 37*(2), 190-203.

Vrij, A. (2000). Detecting lies and deceit: The psychology of lying and implications for professional practice.

Vrij, A., Fisher, R., Mann, S., & Leal, S. (2006). Detecting deception by manipulating cognitive load. *Trends Cogn Sci, 10*(4), 141-142.

Vrij, A., & Ganis, G. (2014). Theories in deception and lie detection. *Credibility assessment: Scientific research and applications*, 301-374.

Vrij, A., Granhag, P. A., Mann, S., & Leal, S. (2011). Outsmarting the liars: Toward a cognitive lie detection approach. *Current Directions in Psychological Science, 20*(1), 28-32.

Vrij, A., & Mann, S. (2001). Telling and detecting lies in a high-stake situation: The case of a convicted murderer. *Applied Cognitive Psychology, 15*(2), 187-203.

Waid, W. M., Orne, E. C., Cook, M. R., & Orne, M. T. (1978). Effects of attention, as indexed by subsequent memory, on electrodermal detection of information. *Journal of Applied Psychology, 63*(6), 728.

Waid, W. M., Orne, E. C., & Orne, M. T. (1981). Selective memory for social information, alertness, and physiological arousal in the detection of deception. *Journal of Applied Psychology, 66*(2), 224.

Walczyk, J. J., Roper, K. S., Seemann, E., & Humphrey, A. M. (2003). Cognitive mechanisms underlying lying to questions: Response time as a cue to deception. *Applied Cognitive Psychology, 17*(7), 755-774.

Walter, W. G., Cooper, R., Aldridge, V., McCallum, W., & Winter, A. (1964). Contingent negative variation: an electric sign of sensori-motor association and expectancy in the human brain. *Nature, 203*(4943), 380-384.

Wang, L., Kuroiwa, Y., Li, M., Wang, J., & Kamitani, T. (2001). Do P1 and N1 evoked by the ERP task reflect primary visual processing in Parkinson's disease? *Documenta ophthalmologica, 102*(1), 83-93.

Ward, T. B., Smith, S. M., & Vaid, J. (1997). Conceptual structures and processes in creative thought.

Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence*: Psychological Corporation.

Wickens, C., Kramer, A., Vanasse, L., & Donchin, E. (1983). Performance of concurrent tasks: a psychophysiological analysis of the reciprocity of information-processing resources. *Science, 221*(4615), 1080-1082.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*(1), 103-128.

Winkler, I., Haufe, S., & Tangermann, M. (2011). Automatic classification of artifactual ICA-components for artifact removal in EEG signals. *Behavioral and Brain Functions, 7*(1), 30.

Winograd, M. R., & Rosenfeld, J. P. (2011). Mock crime application of the Complex Trial Protocol (CTP) P300-based concealed information test. *Psychophysiology, 48*(2), 155-161.

Woodman, G. F. (2010). A brief introduction to the use of event-related potentials in studies of perception and attention. *Attention, Perception, & Psychophysics, 72*(8), 2031-2046.

Wright, G. R., Berry, C. J., & Bird, G. (2012). “You can't kid a kidder”: association between production and detection of deception in an interactive deception task.

Frontiers in human neuroscience, 6.

Wright, G. R., Berry, C. J., Catmur, C., & Bird, G. (2015). Good liars are neither ‘dark’nor self-deceptive. *PloS one, 10*(6), e0127315.

Wu, H., Hu, X., & Fu, G. (2009). Does willingness affect the N2-P3 effect of deceptive and honest responses? *Neuroscience letters, 467*(2), 63-66.

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature methods, 8*(8), 665-670.

Zabelina, D. L., & Robinson, M. D. (2010). Creativity as flexible cognitive control. *Psychology of Aesthetics, Creativity, and the Arts, 4*(3), 136.

Zabelina, D. L., Saporta, A., & Beeman, M. (2015). Flexible or leaky attention in creative people? Distinct patterns of attention for different types of creative thinking. *Mem Cognit*. doi:10.3758/s13421-015-0569-4

Zabelina, D. L., Saporta, A., & Beeman, M. (2016). Flexible or leaky attention in creative people? Distinct patterns of attention for different types of creative thinking. *Mem Cognit, 44*(3), 488-498.

Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics, 13*(1), 75-98.

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. *Advances in experimental social psychology, 14*, 1-59.