**University of Plymouth**

**PEARL** **https://pearl.plymouth.ac.uk**

04 University of Plymouth Research Theses 01 Research Theses Main Collection

2018

# The role of counterfactual thinking in deceptive communication

Briazu, Raluca Andra

http://hdl.handle.net/10026.1/10765

http://dx.doi.org/10.24382/565
University of Plymouth

*All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.*

## Copyright Statement

The copy of this thesis has been supplied on the condition that anyone who consults it is understood to recognize that its copyright rests with its author and that no quotation from the thesis and no information from it may be published without the author's prior consent.

# THE ROLE OF COUNTERFACTUAL THINKING IN DECEPTIVE COMMUNICATION

by

**RALUCA ANDRA BRIAZU**

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

**DOCTOR OF PHILOSOPHY**

January 2018

**Acknowledgements**

Firstly, I would like to express my gratitude to my principal advisor Dr. Clare Walsh for her ongoing encouragement, unremitting patience and her expert advice. Thank you for always having your door open for me, I couldn't have asked for a better supervisor.

Alongside my director of studies, I would like to thank the rest of my supervisory team: Dr. Giorgio Ganis and Dr. Catherine Deeprose, for their insightful comments and encouragement.

My sincere appreciation also goes to Prof. Ruth Byrne for welcoming me so warmly as part of her lab at Trinity College Dublin and for her collaboration in Chapter 4. Huge thanks go to my fellow Cognovian, Chun-Wei Hsu for allowing our collaboration on Study 6, I am so glad we got to work together. I would also like to acknowledge Prof. Sue Denham, for the financial support which enabled my secondment, and for the many unique opportunities for personal and professional development. Also, to Prof. Michael Punt for organizing the 'Topics and concerns' seminars which greatly enriched my postgraduate experience.

To all my fellow Cognovians, it has been a privilege getting to know you, working alongside you and learning from you. A special thank you goes to Katie, Agi and, Diego whose encouragement and friendship have given me strength to continue at many points during the journey. I am also grateful to my brother Sergiu, who has always tried to keep my spirits up. Moreover, thank you to David, my fellow Cognovian and now my husband, I am so pleased I get to share it all with you.

This work is dedicated to Gabriela Briazu, my mother, who has always encouraged my curiosity and kindly financed by endeavours. It is her story-telling that first ignited my imagination and her guidance that first taught me about deception.

## Author's Declaration

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

Study 6 was performed in collaboration with PhD student Chun-Wei Hsu who designed and collected data for the first part of the study. The author of this thesis designed and collected data for the second part of the study. The design of the overall study and the analysis of the data is the author's own work.

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

Studies in Chapter 4 were conducted in collaboration with Prof Ruth Byrne during the author's 3 months secondment at Trinity College Dublin

**Publications:**

Chapter 2 and 3

Briazu R.A., Walsh, C.R., Deeprose C., & Ganis G. (2017). Undoing the past in order to lie in the present: counterfactual thinking and deception. *Cognition, 161,* 66-73.

**Presentations and Posters:**

Briazu R .A. (2016, August). Thinking counterfactually, acting immorally. International Conference on Thinking (ITC), Providence, RI.

Briazu R. A. (2016, June). The role of counterfactual thinking in deceptive communication. European Association for Social Psychology Small meeting on Counterfactual thinking in causality, emotion, communication, and behaviour. Aix-en-Provence, France.

Briazu R. A. (2016, June). What ifs and lies. Poster presented at the 8th Annual School of Psychology Conference. Plymouth University, UK

Briazu R. A. (2016 , January). Counterfactuals and lies. Reasoning and Imagination Lab Talk at Trinity College Dublin. Dublin, Ireland

Briazu R. A. (2015, September) Undoing the past so as to lie in the future – counterfactuals and deception. Poster presented at the European Society for Cognitive Psychology Conference (ESCoP), Paphos, Cyprus.

Briazu R. A. (2015, August). Undoing the Past so as to Lie in the Future – Linking Counterfactuals and Deception. Deception: International conference on Deceptive Communication, Cambridge, UK.

Briazu R. A. **(**2015, June). The role of counterfactual thinking in deception. The 7th Annual School of Psychology Conference. Plymouth University, UK

**Word count for the main body of this thesis:** 42,991

Signed:_____

Date:_____

**The role of counterfactual thinking in deceptive communication**

by

Raluca Andra Briazu

This thesis explores the proposal that there is a close link between counterfactual thinking and lying. Although both require the imagination of alternatives to reality, research has yet to establish a direct link. In the first seven studies the relationship between counterfactuals and lies is directly investigated using novel scenario-based and behavioural tasks. In a further four studies we also investigate the role of affect and executive functions as explanatory mechanisms.

Results show that individuals with a tendency to think counterfactually are more likely to generate potential lies and to be more successful when lying in front of others (Study 1 and 6). Furthermore, we also show that counterfactual availability influences people's tendency to come up with lies (Studies 2, and 3) and the extent to which they expect others to lie (Studies 4, and 5). We also find that the saliency of counterfactual alternatives can affect people's moral standards by motivating them to lie (Study 7). Based on these results we argue that counterfactuals motivate lying by providing information about how things could have been different. We however also investigate alternative explanations. In Studies 8, 9 and 10 we seek to understand whether counterfactually derived affect might also underlie the relationship, but find no such link. Additionally, in Study 11 we investigate the relationship in Parkinson's disease participants in order to understand if executive function might be an underlying mechanism. We do not find this to be the case and we show that PD patients are able to engage in counterfactual thinking and also lie.

The findings in this thesis are the first to provide a direct link between counterfactual thoughts and lies. Overall, we show how counterfactuals can help us mislead others and we reveal that counterfactual thinking is an important cognitive process in deception.

## Ethical Statement

All of the studies in this thesis received ethical approval from the Health and Human Sciences Ethics Committee at Plymouth University. Original application number: 13/14-287 17-794 with amendments: 15/16-489 which allowed the participation of individuals with Parkinson's disease and amendment number 15/16-518 which allowed for performing the studies online.

Due to the nature of the research, in all studies, participants were not fully briefed as to the aim of the research. The studies were advertised as studies about social interaction and following participation, participants were fully debriefed and given the opportunity to withdraw.

## Table of Contents

## List of Appendices

## List of Tables

## List of Figures

# CHAPTER 1

*'Imagination is the deceptive part in man, the mistress of error and falsehood'*

-Blaise Pascal, *Pensées*, 1669

Imagination, our ability to consider things beyond reality, is known to play a crucial role in human cognition (Byrne, 2005). But whilst most consider imagination essential for innovation, some also show that imagination can mislead us. Counterfactual thinking, a central part of imagination, has indeed been shown to lead to erroneous judgements. As the starting quote suggests, counterfactual thinking which represents the mental simulation of alternatives to the past, can lead to false assumptions and errors in judgement (Roese, 1999). However, whilst much research has focused on how counterfactual thoughts can mislead us, it is less clear whether such thoughts can enable us to mislead others. This is surprising given that both counterfactuals and most lies entail a mental simulation of an alternative world in which things could have happened differently (Debey, De Houwer, & Verschuere, 2014; Lewis, 1973). Despite this fundamental similarity, current research only shows how the observation of desired alternatives can encourage us to lie (Shalvi, Dana, Handgraaf, & De Dreu, 2011). As yet no research has examined whether the mental simulation of counterfactuals can have the same influence.

The current thesis aims to address this gap in the literature by describing the first set of studies to directly investigate the relationship between mental simulations of counterfactual thoughts and lying. In doing so, we aim to make a novel

contribution to both our understanding of the functions of imagination and our understanding of the cognitive processes that underline the tendency to lie.

The aim of the current chapter is to review the evidence which points towards a relationship between counterfactual thinking and deception whilst emphasizing the necessity and importance of studying the relationship between these two cognitive processes in a more direct manner. First, we will define the meaning of counterfactual thinking and deception as used within the context of this thesis. Secondly, we will review evidence about factors which influence the generation and content of counterfactual thoughts whilst concomitantly highlighting that some of these factors are also associated with deception. We will also review current knowledge about how such thoughts can be utilized in order to emphasize the way in which our research can contribute to our understanding of counterfactuals. Next, we will emphasize the value of studying the role of counterfactual thinking in deception by evaluating the limited knowledge concerning the cognition underlying the construction of lies. We will also review evidence from separate studies which point towards mechanisms shared by both processes, as well as the few studies which have investigated aspects of counterfactual thinking and deception concurrently. Finally, we will state our research aims and provide an overview of the experimental chapters which follow.

## Definition of terms

Both counterfactual thinking and deception can take many forms, it is therefore important to define the way in which we will use these terms in the current thesis. The definition of terms also allows us to outline the boundaries within which we will be placing the findings of this thesis.

**Definition of counterfactual thinking**

Counterfactual thinking refers to the cognitive process which allows us to consider how past events might have turned out differently (Byrne, 2005; Byrne, 2016; Roese, 1997). As the term 'counterfactual' itself denotes, such thoughts relate to situations that did not actually happen. Counterfactuals are usually characterised by 'what might have been' statements. Most counterfactuals take the form of conditional propositions. These propositions contain an antecedent which corresponds to a state of affairs that isn't true and follows the 'if' preposition and a consequent that corresponds to an outcome and follows the 'then'. For example, thinking 'If I studied harder then I would have passed the exam' constitutes a counterfactual thought. Such thoughts can be used to express a causal understanding of events (e.g. studying leads to improved performance). Crucially, in counterfactual thoughts both the antecedent and the consequent are things that did not happen and are therefore false. Imagination of alternatives to the past can also take the form of semifactuals. These thoughts take the form of 'even if' statements and include a consequent that is true (e.g. even if I had studied harder I would not have passed) (Goodman, 1973). In the current thesis we will focus exclusively on counterfactual thoughts.

Research has established that people create counterfactual alternatives by mentally undoing certain facts in their mental representation of reality (Byrne, 2005; Roese & Morrison, 2009; Roese & Olson, 1993; Watkins, 2008). Most counterfactuals tend to be plausible and believable and tend to be constructed in relation to past events we have personally experienced (Byrne, 2016). However, we are also able to create counterfactuals about things we have never experienced. For example, we can imagine alternatives to past historical events and we can engage in

story-telling by constructing alternative worlds which defy physical reality. In the current thesis we will use the term counterfactual thinking in order to refer to the generation of plausible alternatives to past events that participants can envisage being involved in as these are most pervasive in adult mental life. This is what Schacter, Benoit, De Brigard, and Szpunar (2013) term episodic counterfactual thinking.

**Definition of deception**

Deception is ubiquitous in human interactions, and some argue it constitutes one of the most significant phenomena in human communication (Miller & Stiff, 1993). Although deception is frequently encountered, its definition is not straightforward. Dishonesty can take many forms: for example, one can lie by omitting details or by exaggerating the facts. Alternatively, one can also mislead by telling the truth. As such, there have been many attempts to define the concept of deception and there is not one universally accepted definition. Definitions range from the very broad, according to which the communicator of deception can be anyone, including inanimate beings such as plants (Mitchell, 1986), to more narrow ones according to which deception can occur only between fully responsible and rational individuals (Donagan, 1979). In the context of this thesis however, we will adopt the definition according to which 'deception is intentionally leading another to believe something that the sender knows to be untrue' (Zuckerman, DePaulo, & Rosenthal, 1981, p. 3). This definition is preferred because it integrates two essential elements: intentionality and the sender's belief in the falsity of the information (Masip, Garrido, & Herrero, 2004). Firstly, in order to deceive one needs to generate a false statement with the intention that the recipient believes this to be true. This

means that statements which are untrue but are not meant to mislead, such as jokes or ironic statements, are not lies. Secondly, the sender's belief in the falsity is required in order to make the differentiation between lies and miscommunication. Someone that provides false information because they are not aware of the truth is not considered a liar according to this definition (Knappe & Comaden, 1979). The reverse is also the case, objectively truthful information can constitute a lie if the sender believes that such information is false. This definition implies that we will not be considering double-bluffing or paltering (i.e. lying by telling the truth) given that in such cases the sender knows that information they are conveying is true. This is important as we are interested in how people lie by conveying false information. It is also important to note that, although some also distinguish between deceiving and lying (Stokke, 2013), in this thesis these terms will be used interchangeably. Furthermore, although lies can regard events in the past, as well as intentions for the future, in this thesis we will only concentrate on the former (Vrij, Leal, Mann, & Granhag, 2011).

## Introduction to counterfactual thinking and deception

This section will briefly review the factors which influence the activation and content of counterfactual thoughts as well as the ways in such thoughts can be useful. In doing so our aim is two-fold. Firstly, this evidence will enable us to emphasize the ways in which counterfactual thoughts can be elicited. Such findings constitute the foundation for our experimental manipulation of counterfactual thoughts as presented in future experimental chapters. Secondly, in reviewing current knowledge concerning the uses of counterfactual thinking we can highlight the way in which our research could add to this understanding. Given that our goal is to

emphasize the connections between deception and counterfactuals, in this subsection we will also aim to draw parallels between evidence concerning counterfactual thinking and evidence about deception.

**Activation of counterfactual thoughts and deception**

The factors which determine the spontaneous activation of counterfactual thoughts are usually linked to the characteristics of the outcome to be undone. For example, we are more likely to undo an event which is negative. Negative events such as bereavement can reliably predict the frequency of counterfactuals (Davis, Lehman, Wortman, Silver, & Thompson, 1995). We are also more likely to spontaneously come up with counterfactuals following failure rather than success (Roese & Hur, 1997; Roese & Olson, 1997; Sanna & Turley, 1996). This is driven by negative affect, for example in the study by Davis et al. (1995), the worse one felt immediately after the death of a loved one, the more counterfactuals they came up with months later.

The perceived closeness of an outcome to a more desired, alternative outcome has also been found to influence the generation of counterfactuals. In general, people are more likely to think of alternatives if a desired outcome is narrowly missed either temporally, physically or numerically (Kahneman & Tversky, 1982b). For example, Meyerslevy and Maheswaran (1992) showed that participants engaged in more spontaneous counterfactuals when presented with a scenario in which an actor could have avoided the loss of possession by acquiring insurance 3 days earlier as opposed to when they had the chance to acquire insurance 6 months previously.

In line with the evidence presented above, outcome valence has also been linked to increased deception. For example, one is more likely to lie following a negative event such as the rejection of one's efforts. In a similar manner to counterfactual thoughts, this effect is mediated by negative emotion (Van Der Zee, Anderson, & Poppe, 2016). Additionally, decreased proximity to the desired outcome has also been associated with deception. Individuals tend to lie the most when the difference between the truth and the lie is small in magnitude. For example, it is more likely for participants to lie about the outcome of a die roll if this outcome differs only minimally from the target number which brings financial rewards (i.e. rolling a 5 when the target number for winning a prize is 6) (Hilbig & Hessler, 2013). In a similar manner, academic papers tend to include misreported p-values when the values are only marginally larger than the target value of .05 (Nuijten, Hartgerink, van Assen, Epskamp, & Wicherts, 2016).

**Content of counterfactual thoughts and lies**

The ways in which we undo past events is driven by factors such as norms. Most often we construct counterfactuals in order to reinstate the normal state of events, a finding in line with Norm Theory (Kahneman & Miller, 1986). When presented with a scenario in which an actor ends up having a fatal accident whilst driving home after work, participants tend to undo elements portrayed as unusual. If the actor had the accident on the usual route after leaving early from work, participants tend to think 'if only the actor hadn't left work early'. On the other hand, if the actor left work on time but had the accident whilst driving down a new route, participants tend to think 'if only the actor had driven the usual way'. These findings have been widely supported (Buck & Miller, 1994; Lundberg & Frost, 1992; Wells,

Taylor, & Turtle, 1987) and suggest that we tend to construct counterfactuals in order to alter exceptional events so that they agree with normality.

Kahneman and Tversky (1982b) also showed that actions are easier to undo than inactions, and this could be because we perceive inactions as normal and actions as abnormal (Kahneman & Miller, 1986). A person who had lost a sum of money following the decision to move shares from one company to another will be judged as more likely to experience counterfactually derived emotions such as regret, in comparison to a person that experienced the same loss following the decision to not switch companies (Kahneman & Tversky, 1982a). That is, it is easier to imagine how an event could have been undone than imagine an action that was never performed.

We also find it easier to undo controllable antecedents rather than antecedents we perceive having little control over. Girotto, Legrenzi, and Rizzo (1991) presented participants with a scenario that contained several incidents which prevented the protagonist from getting home on time. Results showed that participants were more likely to mutate aspects that were in the protagonist's control rather than ones that were uncontrollable. For example, participants were more likely to think counterfactually about the protagonist's decision to stop for a beer, as opposed to the delay caused by the flock of sheep that crossed the road. It is worth stating that this effect can be modulated by other factors, such as the perceived appropriateness or inappropriateness of the events (McCloy & Byrne, 2000) and whether such events are observed or actually experienced (Girotto, Ferrante, Pighin, & Gonzalez, 2007). Moreover, Mercier et al. (2016) questioned the robustness of this effect by showing that episodic counterfactual thought tends to focus mostly on

uncontrollable aspects unless people are prompted otherwise (but see Roese, Smallman, & Epstude, 2017).

Although evidence concerning the content of deceptive communication is scarce, there is some evidence to suggest that norms and control are also linked to unethical behaviour. For example, behaviour which violates norms is likely to be judged as more dishonest (Levine et al., 2000). That is, when behaviour is regarded as socially inappropriate we tend to attribute it lower credibility. Therefore, we are more likely to consider both counterfactuals and dishonesty in relation to unusual events.

In line with the controllability effect, but from a slightly different vantage point, control has also been associated with deception. Locus of control, that is the extent to which one feels life events are controlled by oneself, has been found to influence ethical behaviour (Trevino, 1986). Specifically, it has been shown that individuals with an internal locus of control are more likely to be honest (Hume & Smith, 2006). This is because a person who tends to see themselves responsible for their outcomes will be more likely hold themselves accountable for the consequences of an unethical behaviour. Conversely, people with an external locus of control tend to be less ethical because they do not believe in the connection between their actions and their misdeeds. If, as Mercier et al. (2016) suggest, spontaneous counterfactuals in response to a negative event are mostly about uncontrollable factors, then counterfactuals might provide the necessary material for individuals to disassociate from their actions thus enabling them to be unethical and not take responsibility for their misdeeds.

**Uses of counterfactual thoughts**

The mind computes counterfactuals for several reasons. Counterfactuals can help us justify and excuse the past, modulate our emotional experiences and allow us to prepare for the future (for a review see Byrne, 2016). Thinking about how things outside of our control could have been different can excuse poor performance and help us explain past events. For example, one can excuse poor performance in an exam by thinking 'If only there was more time' (McCrea, 2008; Tyser, McCrea, & Knüpfer, 2012). Similarly, reflecting on how things could have been worse by engaging in downward counterfactual thoughts can elicit contrastive and assimilative effects which allow us to justify past behaviours. For example, Markman, Mizoguchi, and McMullen (2008) showed that statements about the torture of prisoners in Iraq which focused on how things could have been worse helped participants feel less negative and further lowered ethical standards regarding the treatment of prisoners of war in the future. Catellani and Bertolotti (2014) have also shown that politicians who excuse their actions by focusing on what others could have done differently tend to be judged as less responsible for the outcome as well as more believable and convincing. This implies that counterfactual thoughts can constitute an effective defence strategy. This evidence is relevant in connection to deception as it potentially suggests that counterfactual thoughts could provide relevant material which can be used to misguide and manipulate others.

We also construct counterfactuals in order to regulate our emotions. Certain counterfactuals are known to amplify negative emotions such as regret, guilt and shame (Zeelenberg et al., 1998). However, such thoughts can also help dampen negative emotional experiences. Survivors of natural disasters who concentrate on

how things could have been worse regard themselves as lucky survivors rather than unlucky victims (Teigen & Jensen, 2010). Sweeny and Vohs (2012) also showed that participants experienced relief following the near miss of a terrible outcome if they engaged in downward counterfactual thinking. This suggests that some counterfactual thoughts can be used to make us feel better.

The need to regulate our emotion can also elicit deception. Lies can be used as a defensive tactic to protect positive self-views (Feldman, Forrest, & Happ, 2002). For example, following a wrong-doing, the more blameworthy a participant feels the more likely they are to lie. This suggests that sometimes we lie in order to avoid feelings of guilt. Furthermore, there is also evidence to suggest lies are also told in order to avoid embarrassment and disapproval (DePaulo, Kirkendol, Kashy, Wyer, & Epstein, 1996). Therefore, just like certain types of counterfactual thoughts, lying can help us avoid negative emotions and make us feel better about what we have done in the past.

For a long time, the primary role of counterfactuals was thought to be related to its preparatory function. A persuasive account for this argument is the Functional Theory by Epstude and Roese (2008). The theory is underpinned by the idea that counterfactuals direct behavioural intentions through a content specific pathway or a content-neutral one. The content specific pathway specifies that information contained within the counterfactual thought can directly influence future behaviour. For example, experiencing a negative event (i.e. being sunburned), enables the activation of a counterfactual thought (i.e. I should have worn sunscreen) which later lead to a behavioural intention (i.e. now I will wear sunscreen) (Smallman & Roese, 2009).

Conversly, the content neutral pathway points out that counterfactuals can influence behavior irrespective of the specific information they contain, by eliciting mind-set priming, contrast effects, and motivations. In line with this is evidence suggesting that engaging in a counterfactual thinking mind-set helps performance on an unrelated subsequent task. For example, thinking about what could have been by removing antecedent elements can enable a relational processing style which can increase performance on a subsequent unrelated analytical task (Kray, Galinsky, & Wong, 2006). The content-neutral pathway also works by way of motivational effects. Markman, McMullen, and Elizaga (2008) showed for example how failure can provoke negative affect which subsequently can elicit greater effort on a consequent task. In the study, participants who generated counterfactuals about their performance on an anagram task, went on to perform better on a subsequent similar task.

The functional nature of counterfactual thoughts has however been partially challenged. New findings suggest that individuals who have just failed a task often modify uncontrollable features of past actions when asked to think how things could have been better (Mercier et al., 2016). Uncontrollable features do not provide information for how behaviour can be modified in the future, and so this provides a challenge for the preparatory function of counterfactuals. This suggests that counterfactual thinking might have other uses. As shown above lying is activated by similar factors and for similar reasons as counterfactuals, as such we would like to suggest that counterfactual thinking might be an important cognitive process for lying. The investigation of this hypothesis is important as currently our understanding of the cognitive processes underlining the generation of lies is limited. In the following section we will review our current knowledge of the cognitive

processes that contribute to the construction of lies whilst aiming to highlight the potential value of identifying counterfactual thinking as one of the cognitive processes involved in deception.

## Cognitive perspective on deception

Due to its wide-spread use, and its potential harmful consequences (DePaulo, Ansfield, Kirkendol, & Boden, 2004), research has mainly focused on the effective ways to detect lies (Gamer & Ambach, 2014). In contrast, the cognition of deception has largely been overlooked. This is surprising given that in order to better detect deception we need to first understand how lies come to be. Walczyk, Harris, Duck, and Mulay (2014), argue that although a normative account for deception detection requires a strong understanding of the underlying mechanisms, current methods are developed based on weak theoretical foundations that lack rigorous scientific testing. Levine and McCornack (2014) postulate that in order to improve deception theory there needs to be more emphasis on the way in which deception discourse is produced. It is thought that an understanding of how deception is formulated would further theoretical predictions which in turn could enable researchers to detect deception using high levels of accuracy. Despite these recommendations, identifying the cognitive processes which contribute to deception, modelling their interaction and their relation to personality and situational factors remains one of the major challenges of deception research (Gamer & Ambach, 2014).

Whilst theories have evolved beyond simply emphasising that deception is multifaceted in nature (Zuckerman et al., 1981), we still are some way off identifying all of the mechanisms that contribute to deception (Gamer & Ambach, 2014).

Currently, one of the most coherent theoretical accounts of how we come up with lies is given by the construction component of the Activation-Decision-Construction Action Theory (ADCAT) (Walczyk et al., 2014) [1]. According to the theory, the construction of a lie is based on the activation of semantic and episodic memories relevant to the topic an individual aims to lie about. The activation of relevant memory nodes will help bring the truth into working memory together with other interrelated information that could be used during the later construction phase. If a person decides to lie they need to refrain from producing the truth which is also active in working memory and fabricate the lie by relying on associated sematic and episodic nodes. The theory also emphasizes that the construction of a lie is constrained by other factors such as what the target of the deception might already know. The theory thus highlights that lying is a cognitively effortful process because it requires creating and relating a novel experience that has not actually occurred. In line with this account, evidence suggests that lying is associated with increased response latencies (but see Burgoon, 2015; Verschuere & De Houwer, 2011). Empirical studies show that unrehearsed lies are longer to come up with in comparison to rehearsed lies (DePaulo et al., 2003; Walczyk, Mahoney, Doverspike, & Griffith-Ross, 2009; Zuckerman et al., 1981), and that similarly, lies in response to open ended questions are slower compared to response to yes/no questions (Walczyk et al., 2005; Williams, Bott, Patrick, & Lewis, 2013), as are lies in response to questions which require longer answers (Vrij, Mann, Kristen, & Fisher, 2007).

[1] Other cognitive accounts for deception exist. For example the Working Memory Model (Sporer & Schwandt, 2007), the Cognitive load paradigm (Vrij & Granhag, 2012) and the Interpersonal Deception Theory (Buller & Burgoon, 1996) have all been influential for the field. We are not reviewing these theories here as the ADCAT combined insights from all these models whist also providing inferences for how cognitive processes might influence the content of lies.

Beyond showing that constructing lies can impose cognitive load the theory also makes suggestions regarding the type of information we might rely on in order to construct lies. Namely, the involvement of episodic memory in lying implies that we rely on our past experiences to come up with lies. Indeed, empirical research shows that people report lying by slightly changing details of experiences that have already occurred and by relying on the truth (Debey et al., 2014; Malone, Adams, Anderson, Ansfield, & DePaulo, 1997). This suggest that in order to formulate a plausible and socially appropriate lie one needs to think about alternatives to what actually happened. Counterfactual thinking is particularly relevant to this argument. As reviewed earlier counterfactual thoughts are instantiations of events that did not actually happen and they are constructed by minimally altering details relating to past experiences (Byrne, 2016). Just like deception, counterfactual thinking is also underpinned by episodic memory and its underlying neural mechanisms (Schacter et al., 2013). It then follows that counterfactual thinking might be a cognitive mechanism related to lying.

To summarize in this section we have attempted to highlight that there is a need for the identification of cognitive processes that contribute to lying and that counterfactual thoughts which are constructed in a similar manner could be such a process. Yet, the link between these two processes has mainly only been alluded to. In the next sections we will review this evidence.

## Converging evidence for the role of counterfactual thinking in deception

Currently the link between counterfactual thinking and deception can only be inferred. This is because most of the research related to a possible association between the two processes comes from studies which have separately assessed these

variables. The following section will review these studies and will include evidence from developmental, neuroscience and individual differences research. Additionally, we will also review the very few studies which have assessed deception and counterfactuals concomitantly.

**Developmental research**

The capacity to imagine alternative worlds and the ability to deceive appear to follow a similar developmental trajectory. Separate studies from each areas show that children as young as 2 can envisage imaginary situations (Harris, German, & Mills, 1996) and can deny an action they have committed (Evans & Lee, 2013). However, at this age infants make mistakes which means these abilities are far from being adult-like. Whilst 2 year old can think outside the here and now, they find it more difficult to compare and contrast reality with 'what might have been' (Beck & Guthrie, 2011). Similarly, although 2 year olds can deny something they have done they fail to follow through with the lie when asked additional questions (Evans & Lee, 2013).

As age increases children become more skilled in these abilities and in both cases this enhanced ability is linked to the acquisition of theory of mind (ToM). ToM represents our ability to understand that other people can have mental states such as beliefs, desires and knowledge that differ from our own (Goldman, 2012). Once this ability is acquired, around the age of 4 years, children can more easily misinform and withhold information (Peskin, 1992; Wellman, Cross, & Watson, 2001). ToM has in fact been shown to be crucial in the ability to misinform. Ding, Wellman, Wang, Fu, and Lee (2015) have demonstrate its causal role in deception by showing that even children as young as 3 can lie if they participate in ToM training.

Understanding that others can hold beliefs that differ from our own is also associated with counterfactual thinking. Guajardo and Turley-Ames (2004) have shown that participants' performance on false-belief tasks, a frequently used methodology to examine ToM, is positively correlated with counterfactual thinking performance. Riggs, Peterson, Robinson, and Mitchell (1998) also found that in order to respond correctly to some false-belief tasks children need to be able to handle counterfactual situations. In order to accommodate the perspective of another person one needs to modify one's own knowledge of the situations and simulate an alternative reality. Take for example, the popular false-belief task entitled Sally-Ann, where Sally places a marble in a basket in the presence of Ann and then leaves the room during which time Ann removes the marble from the basket and places it within a box instead. Riggs et al. (1998) argued that in order to respond correctly the question asking where Sally will look for the marble upon her return, children need to first think what would have happened had Ann not moved the marble. Children thus need to make a counterfactual inference in order to solve false-belief task. In a similar vein, others have also shown that children find it easier to engage in counterfactual reasoning about mental states than false-belief inferences about such states (Perner, Sprung, & Steinkogler, 2004; Rasga, Quelhas, & Byrne, 2016).

Given that ToM is necessary for deception and counterfactuals are a key ingredient to reasoning in false belief tasks it could be suggested that the ability to deceive also relies on the development of counterfactual thinking ability. Indeed, Evans, Xu, and Lee (2011) have suggested that the sophistication of children's lies could be related to their ability to carry out counterfactual thinking about long causal chains. This hypothesis needs to be considered with caution given that Beck, Riggs, and Gorniak (2010) have shown that children did not perform differently when

considering counterfactuals involving long causal chains. Nonetheless, these lines of evidence do show that the ability to engage in counterfactual thinking might be a pre-requisite to the development of the ability to lie and thus it is surprising that so far no studies have mapped the development of deceptive ability in relation to counterfactual thinking.

Further support for the hypothesis that deception and counterfactual thinking might be associated comes from evidence suggesting that both processes rely on the development of executive functions. Executive functions are processes that govern goal-directed thought and behaviour. Miyake et al. (2000) propose that the processes that contribute to executive functions are independent but interrelated and they include inhibitory control, updating information in working memory and attention shifting. Each of these processes is called for when thinking about counterfactuals and when lying. Inhibitory control is the capacity to suppress a response or behaviour and is needed when lying because one needs to inhibit responding with the truth whilst reporting falsehoods (Talwar & Lee, 2008). Similarly, in order to respond to questions about a counterfactual world one needs to resist from responding with what actually happened (Beck, Riggs, & Gorniak, 2009).

Working memory, a temporary system for holding information in the mind is required when lying because one needs to keep in mind both the lie and the true state of affairs. Alloway, McCallum, Alloway, and Hoicka (2015), for example found that good liars have higher working memory scores. In a similar vein, counterfactual statements require one to acknowledge multiple possibilities and keep in mind both the reality and the imagined alternative (Byrne, 2005). Indeed, children with higher

working memory skills are capable of answering more counterfactual questions correctly (Drayton, Turley-Ames, & Guajardo, 2011; Guajardo, Parker, & Turley-Ames, 2009) although it is also worth mentioning that studies using different methodology did not find the generation of counterfactuals to rely on working memory (Beck et al., 2009).

Attention shifting which refers to our ability to switch between mental sets or tasks is especially needed in the generation of counterfactual emotions such as regret. To experience such an emotional reaction one needs to keep the actual and the counterfactual world in mind and compare them by switching between the two worlds back and forth (Burns, Riggs, & Beck, 2012). Similarly, constructing a plausible lie involves coming up with a false statement whilst making sure deceptive communication is congruent with what the receiver already knows, a process which requires switching between the truth and the lie (Debey, Liefooghe, De Houwer, & Verschuere, 2015). Thus, it is widely agreed that all executive function skills as defined by Miyake et al. (2000) are highly intertwined with both the generation of lies and that of counterfactuals.

Rapid improvements in all of these executive function skills occur during pre-school years and early years of school (Zelazo et al., 2003), which is why the widely held view has been that most of these skills are fully acquired early on. However, recent research in older participants has found that executive functions follow an incremental developmental pattern which extends into adolescence (Best & Miller, 2010). Such changes are both quantitative and qualitative in nature and children become both quicker and more accurate on tests of executive function as they grow older.

Closely linked to the development of executive processes, recent studies have also found that the ability to engage in deception and counterfactual reasoning develops beyond early childhood. Counterfactual thinking improves throughout middle childhood (Beck, Robinson, Carroll, & Apperly, 2006), and Rafetseder, Schwitalla, and Perner (2013) have shown that it is not until the age of 12 that children acquire adult-like counterfactual thinking ability.

The evidence in relation to the development of deception beyond pre-school age is mixed. Although most studies agree that the sophistication of lies is related to improvements in executive function skills, they disagree in regard to how this is related to age. For example, when testing participants aged between 8 and 16 years old, Evans and Lee (2011) found no relationship between age and frequency of deception. Conversely, other studies found age-related differences in deception that generally fit with the U-shaped pattern of age-related changes observed in inhibitory control (Debey, De Schryver, Logan, Suchotzki, & Verschuere, 2015). In particular, Debey, De Schryver, et al. (2015) found lying frequency to peak in adolescence, specifically between the ages of 13 and 17. Thus, this evidence might indicate that one is more likely to engage in deception after adult-like counterfactual thinking is achieved.

Overall, this section has highlighted that the development of both counterfactual thinking and the ability to deceive appear to be closely linked. In the next section we will provide further evidence for the association between these two processes by looking at how they rely on similar neural mechanisms.

**Neuroscience research**

Given the evidence showing that counterfactual thinking and deception are related to other processes such as ToM and executive function it is not surprising that there is no single neural network linked to either deception or counterfactual thinking. Both processes rely on multiple neural networks that cut across different psychological domains. In both cases, the networks that become activated depend on the content of the counterfactual thought, and the type of lie (Ganis, Kosslyn, Stose, Thompson, & Yurgelun-Todd, 2003; Van Hoeck, Watson, & Barbey, 2015). That being said, there is evidence to suggest that on the whole, both processes activate similar neural networks. For example, both processes activate neural networks responsible for internal directed cognition (Abe, 2011; Spreng, Mar, & Kim, 2009) as well as areas linked to autobiographical and episodic memory (Abe et al., 2008; De Brigard, Addis, Ford, Schacter, & Giovanello, 2013; Ganis et al., 2003; Van Hoeck et al., 2014). Generation of both lies and counterfactuals also activate the areas in the medial prefrontal cortex which are known to facilitate mental simulations by integrating memory information and sensory information in the representation of behaviour (Barbey, Krueger, & Grafman, 2009; Yokota et al., 2013).

Given the developmental evidence reviewed earlier in this chapter, it is not surprising that studies also suggest that both processes activate networks responsible for executive function such as the fronto-parietal network (Abe, 2009; Christ, Van Essen, Watson, Brubaker, & McDermott, 2009; Coricelli et al., 2005; De Brigard et al., 2013; Gombos, 2006). Damage to brain regions within this network such as the pre-frontal cortex has also been linked to a decrease in the activation of spontaneous

counterfactual thoughts (Beldarrain, Garcia-Monco, Astigarraga, Gonzalez, & Grafman, 2005). Furthermore, the direct stimulation of the same brain region through the use of transcranial direct current stimulation has also been shown to impair lying ability (Mameli et al., 2010; Priori et al., 2008). Just as the development of counterfactual and deceptive ability follow the maturation of the prefrontal cortex and hence that of executive skills, the decline in such skills which occurs with an increase in age (Glisky, 2007) also leads to impairment in these processes. As we grow older we become slower and less accurate when lying (Debey, De Schryver, et al., 2015). Similarly, in comparison to their younger counterparts, healthy older adults tend to be less likely to engage in spontaneous counterfactual thought or to make counterfactual inferences (Walsh, Deeprose, & Briazu, under review).

Findings of studies with patients known to be impaired in executive function reveal a similar story. Parkinson's disease (PD), a neurodegenerative disorder associated with substantial decrease in dopamine levels in the basal ganglia, frontal lobes and hippocampus has also been associated with a decrease in the ability to engage in counterfactual thinking and lying. PD patients find it more difficult than healthy controls to lie about recognizing a stimulus when instructed to do so. Likewise, PD patients demonstrate decreased performance on a test of counterfactual inference and tend to think less about how thinks could have been better (Abe et al., 2009; McNamara, Durso, Brown, & Lynch, 2003).

Studies in others who exhibit prefrontal dysfunctions, such as people with schizophrenia also show impairment in the generation of counterfactuals relative to heathy controls (Contreras et al., 2016; Hooker, Roese, & Park, 2000). In a related manner, participants with schizotypal personality traits have also been found to be

impaired in their ability to deceive (Barnacz, Johnson, Constantino, & Keenan, 2004).

Overall, the cognitive processes of interest to this thesis seem to rely on similar neural mechanisms and are similarly impaired in populations with damage to the brain areas which are associated with such networks. Despite the fact that people rely on the same mechanisms in order to activate both lies and counterfactuals they tend to differ in their tendency to so. In the next section we will review some individual differences which characterise individuals with an increased tendency to engage in both processes.

**Individual differences research**

People differ in the degree they engage in counterfactual thinking, and likewise some people lie more than others. Similar personality traits have been linked to the tendency to deceive and the tendency to imagine alternative possibilities. Personality traits identifiable within the Five Factor model, such as neuroticism, openness to experience, extraversion and sub-facets of the conscientiousness construct, can reliably predict the generation of counterfactual thoughts (Allen, Greenlees, & Jones, 2014; Bacon, Walsh, & Briazu, under review). All of these personality traits have also been associated with deceptive self-presentation on social media (Michikyan, Subrahmanyam, & Dennis, 2014; Möller, 2016) and with an increased tendency to lie during social interactions (Sarzyńska et al., 2017). The generation of counterfactuals was also found to be negatively related to agreeableness, one of the Big 5 personality traits as measured by the NEO Personality Inventory (Bacon et al., under review), a social trait which reflects individual differences in cooperation and social harmony. Interestingly, individuals

who score highly on the agreeableness component tend to also regard themselves as less successful liars (Giammarco, Atkinson, Baughman, Veselka, & Vernon, 2013).

Both processes have also been associated with creativity. Markman, Lindberg, Kray, and Galinsky (2007) showed that undoing reality by thinking of additional elements that could have improved an outcome (i.e. additive counterfactuals) can promote a mind-set which in turn can lead to better performance on a creative generation task. Creative individuals in turn have been shown to have a lower level of integrity (Beaussart, Andrews, & Kaufman, 2013), and a greater tendency to engage in deceptive behaviour (Gino & Ariely, 2012). Furthermore, Gino and Ariely (2012) also suggest that encouraging creative mind-sets can lead to deception. The positive association between creativity and deception has however been challenged, as Niepel, Mustafić, Greiff, and Roberts (2015) found that although performance on measures of creativity and ethical decision making positively correlate, creativity does not predict ethicality when assessed longitudinally. However as opposed to the initial study this study only used self-report measures of creativity rather than behavioural tasks.

Another individual difference characteristic that links the two processes is belief in a just world. People differ in the degree to which they believe that the world is fair and that one always receives what they deserve (Lerner, 1980). Kasimatis and Wells (1995) mention that those who hold this belief should engage less in counterfactual thoughts as believing that all outcomes are deserved should make the imagination of possible alternatives difficult. Maybe not coincidently, people that believe in a just world are also more likely to be honest (Dalbert, 2009), whilst those

who view society as less meritocratic are more likely to justify tax evasion and bribes (Kline, Galeotti, & Orsini, 2014).

Overall this evidence indirectly suggests that individuals with a tendency to consider counterfactual alternatives might also have a tendency to lie. This is relevant in the context of this thesis as the tendency to lie maps onto the successful ability to do so. For example, Debey, De Schryver, and colleagues (2015) found that individuals who are better at lying do so more frequently. Similarly, older adults' diminished tendency to lie is associated with reduced inhibitory ability, a skill known to contribute to deception success (Haj & Antoine, 2018). Therefore, if those who are successful at lying tend to lie more, and if ability to deceive is a skill that requires counterfactual thinking, then individuals who lie more should have an increased tendency of engaging in counterfactual thinking. However, the link between counterfactual thinking and lying has never been investigated in relation to individual differences. In fact, only a hand-full of studies have investigated counterfactuals and deception concurrently. In the next section we will review these studies.

**Current research on counterfactuals and deception**

One of the earliest studies investigating counterfactuals and deception concomitantly was a study by Miller, Visser, and Staub (2005) which looked at the way in which the imagination of alternative possibilities to another person's behaviour can influence our perception of dishonesty. Miller et al. (2005) found that observing a protagonist resisting the temptation to cheat under surveillance leads to inferences about how he would have cheated had he not been surveyed, and consequently to the perception that such a protagonist is more dishonest than

average. This evidence points towards a potential positive association between counterfactuals and deception and aligns with evidence we have previously reviewed in this chapter.

Alternatively some have also shown that counterfactual thinking can discourage immoral behaviour. Gaspar, Seabright, Reynolds, and Yam (2015) showed that individuals asked to reflect how they could have done things differently in regard to a prior misdeed feel more guilty about their past actions and in turn tend to be less dishonest in a subsequent unrelated task. This suggests that moral affect derived from counterfactual reflections encourages one to morally compensate in a subsequent task. However, in this study the decision to deceive was analysed in the context of prior immoral behaviour, and the results could reflect moral cleansing (West & Zhong, 2015). Moral cleansing occurs when one deviates from what should be and subsequently tries to compensate for such behaviour by engaging in behaviour which contradicts their prior actions. Thus, thinking about how one could have avoided their prior misdeed can increase the threat to the self-image, which in turn could trigger a need to compensate by engaging in morally correct behaviour. Conversely, counterfactual thinking of how one's action could have been worse has been shown to increase feelings of security in one's moral self-imagine which in turn can allow participants to feel licensed to act in morally dubious ways (Effron, Miller, & Monin, 2012). For example, if individuals can easily point towards an instance in the past where they avoided exhibiting racist behaviour, they subsequently tend to be less racially sensitive. These findings provide interesting evidence as to the way in which counterfactual thinking can impact moral behaviour. However, they do not allow us to understand whether counterfactual thinking about an event can influence subsequent deception about that same event. This is because in these studies the

generation of counterfactuals focused on the misdeed rather than the events which lead to the misdeed. Also in the study by Gaspar et al. (2015) the counterfactuals generated were about an event that was not directly related to the subsequent immoral behaviour.

As far as we are aware, only one study assessed the impact of counterfactual availability on subsequent deception. Shalvi et al. (2011) studied the effect of observing desired alternatives to a die roll on subsequent lies about the outcome of that die roll. During the study participant had to roll a die under a cup, this was so that the outcome of the die was private. Participants were told to either roll the die once, or roll it three times. In each condition participants were asked to only report the outcome of the first roll. Participants were told that they would be paid according to the outcome they reported, i.e. the higher the reported outcome, the higher the pay-out. Results showed that participants who could roll the die more than once lied more. The authors reasoned that these participants had more of an opportunity to observe desired alternatives (i.e. higher outcomes than the one initially rolled), and thus the observation of something that could have happened allowed them to justify lying about the outcome. In follow up studies the authors also showed that lying following the observation of a counterfactual tends to be regarded as more morally acceptable. Shalvi et al. (2011) have therefore suggested that counterfactuals motivate deception because they enable self-justifications which allow the individual to profit from lying whilst feeling honest. Such findings have been replicated by Bassarak et al. (2017) who have additionally found that observed counterfactuals can influence deception even when the outcome of the die roll is determined by someone else. This evidence suggests that it is the information contained in the

counterfactuals that influence deception and not the characteristics of the situations itself such as whether or not one had ownership over the die roll.

One limitation of using the die roll paradigm also known as the die-under-cup task is that both counterfactual thinking and deception had to be inferred rather than directly measured. Because the die rolls were done in private to assure participants' decisions to lie would not be influenced by the possibility of being caught, deception had to be inferred by comparing reported outcomes against a standard distribution of die rolls. Thus, this paradigm prevents the identification of those who actually lied, as individual reports cannot be traced, and data needs to be analysed at the group level.

Furthermore, this paradigm does not allow for the direct assessment of counterfactual thoughts. Even though Shalvi et al. (2011) specifically focus on the observation of desired counterfactuals, that is an observation of higher die rolls which are linked to a higher pay off, the participants could have also observed worse counterfactuals. Therefore, it is not clear whether only certain types of counterfactuals influence deception or whether deception is influenced by consideration of how things could have been different in general. Another important point in relation to the assessment of counterfactuals in this study is that the observation of additional die rolls might not require the mental simulation of alternatives in the same way that spontaneous counterfactuals do. This is important given that most counterfactual thoughts occur automatically through the mental stimulation of alternatives (Byrne, 2005). Given these limitations, the question remains, is the ability to engage in counterfactual thinking related to the ability to deceive? The current thesis aims to address this question by directly investigating the

role of counterfactual thinking in deceptive communication. This will be the first thorough investigation aimed to understand the role of counterfactual thinking in deceptive communication.

## Aims and thesis structure

Based on the evidence presented in this introductory chapter we hypothesise that counterfactual thinking is an important process linked to our tendency and ability to deceive. In order to assess this hypothesis the current thesis aims to investigate whether counterfactual thinking and deception are associated whilst also examining the mechanisms that might underlie this relationship. The thesis is organised in two parts. In the first two experimental chapters that follow we aim to investigate the associations between counterfactual thinking and lying both in terms of individual differences and though the experimental manipulation of counterfactual information and assessment of its impact on subsequent deception. In Chapter 2 we aim to do so by utilizing scenario-based measures. In Chapter 3 we assess the role of counterfactual thinking in deception by using paradigms which more closely resemble real-life situations. The second part of this thesis is dedicated to the investigation of potential underlying mechanisms. In Chapter 4 we aim to investigate the whether affect derived from counterfactual thought might impact the relationship between counterfactual and deception. In Chapter 5 we focus instead on executive functions in order to understand whether the relationship between counterfactuals and lies could in part be explained by a shared reliance on these processes. Finally, in Chapter 6 we provide a summary of findings and highlight potential implications for both counterfactual thinking and deception.

# CHAPTER 2

## Are counterfactual thinking and deception associated?

As reviewed in Chapter 1, there is some evidence to suggest that counterfactual thinking could be linked to deception. Evidence from developmental, neuroscience and individual differences studies all point towards an association between these two processes. However, only a limited number of studies have attempted to examine these processes concurrently and a direct link has yet to be established. The aim of the current chapter is to empirically explore the association between counterfactual thinking and deceptive statements.

As of yet only Shalvi et al. (2011) have investigated the effect of counterfactuals on deception. The study showed that the observation of desirable outcomes can bring to mind events that almost happened, which allows participants to justify their deception. However, as discussed earlier, individuals only observed alternatives rather than mentally simulating them. Therefore, the question remains, are counterfactual thoughts and deception associated? If so, how does the simulation of alternatives to past events influence subsequent deceptive communication? The present chapter aims to answer these questions by examining the link between counterfactuals and lies using scenario-based measures which allow for the direct assessment of both processes. Scenarios have extensively been used in the study of counterfactual thinking (Byrne, 2005), and are also preferred in the study of deception as this can be a sensitive topic to assess directly (Van Der Zee et al., 2016). Whilst hypothetical scenarios can differ from real-world situations in many

regards, such measures have been shown to provide responses which are sufficiently similar to real-life experiences (Hughes, 1998).

In this chapter scenario-based measures were designed by manipulating factors known to stimulate mental representations of counterfactuals. Our aim was to assess whether the predisposition to engage in counterfactual thinking is associated with the propensity to deceive (Study 1) and to examine whether changing the availability of counterfactual alternatives impacts individual's subsequent deceptive responses (Study 2 and 3) and inferences about the likelihood that someone will lie (Study 4 and 5). Overall the current chapter aims to clarify the relationship between imagining alternatives to the past and deceptive communication.

## Study 1

Our first study focused on the relationship between counterfactuals and deception using an individual differences approach. People differ in the ability and degree to which they engage in counterfactual thought (Allen et al., 2014) and deception (Sarzyńska et al., 2017). As counterfactual thinking and deception have not been assessed simultaneously, we wanted to assess whether these tendencies are related within a single study. We tested people's spontaneous tendency to produce counterfactual thoughts and deception, and also their ability to generate these when cued. If counterfactual thinking and deception share the same underlying processes, then we should find a positive association between the spontaneous tendency to think counterfactually and lie and between the ability to generate counterfactuals and lies when cued. Spontaneous and cued generations are governed by different

mechanisms (Beldarrain et al., 2005) therefore we do not expect these to be associated.

Additionally, given that anxiety and executive function have been closely linked to both deception and counterfactual thinking (Prokopcáková & Ruiselová, 2008), we also wanted to investigate whether the relationship between counterfactuals and lies could be explained by a shared association with these variables.

## Method

### Participants

The participants were 81 undergraduate students at Plymouth University who received course credit for their participation. One participant did not complete all the tasks and was therefore excluded from the analysis. The remaining participants (60 female and 20 male) ranged in age from 18 to 40 years ($M$ = 20.23, $SD$ = 3.20).

### Materials and Procedure

All participants completed a battery of tests which included two scenarios to assess counterfactual thinking and deception, four measures of executive function and a questionnaire to assess anxiety. The measures are described in detail below.

**Scenarios.** The scenarios included here were chosen based on a pilot study which assessed participants' performance on four different scenarios. We chose the scenarios which provided the widest range of responses in terms of counterfactuals and deception. The two scenarios included were developed based on factors known to affect mutability of events (Byrne, 2002; Kahneman & Miller, 1986; Roese & Olson, 1995). Counterfactuals are more likely to follow negative events (Roese,

1997), therefore each scenario described events during which participants were asked to imagine making a series of decisions which lead to an unexpected negative outcome. Each scenario was followed by questions probing whether participants would be inclined to lie about elements included in the scenarios. As deceptive responses depend on the potential outcome for the liar (ten Brinke & Porter, 2012), one of the scenarios involved a low risk opportunity for deception (lying to a neighbour) whilst the other involved a higher risk (i.e. lying to the police).

The scenario which involved a high risk was the 'moving town' scenario which was adapted from McEleney and Byrne (2006). In short, for this scenario participants had to imagine moving to a new town and making decisions which result in difficulties meeting new friends. These decisions include: moving town, going to a movie rather than a neighbour's party and buying a new stereo rather than joining the gym with a work colleague.

The scenario which involved a low risk was the 'car incident' scenario which developed specifically for this study. In this case participants were asked to imagine making a series of decisions which lead to a minor car incident. These included going shopping to a supermarket rather than the corner shop, responding to a phone call whilst driving instead of ignoring it and driving down a new route rather than a more familiar one. A full version of each scenario as presented to the participants is included in Appendix 2A.

Participants were given a 12 page booklet and responded to all questions in writing. They first received the scenarios in a counterbalanced order but always answered counterfactual and deception questions in a similar order as explained below. First, *spontaneous counterfactual thinking* was elicited by giving them 5

minutes to write a diary page about their imagined experience. After completing both diary pages, we asked questions assessing dishonesty. To measure *spontaneous deception,* participants received an indirect and a direct question for each scenario. For the 'moving town' scenario the indirect question asked participants to write down anything they would say to their neighbour when meeting them after the party they had failed to attend. The direct question asked them to write any specific reason they would give for not attending the party. For the 'car incident' scenario, the indirect question asked participants to write down anything they would say to the police if they were to come to their door. The direct question asked them to write down anything they would say to the police when specifically questioned on their knowledge concerning the damaged car. No time limit was imposed for answering these questions. *Cued deception* was measured by asking participants to write down all the things they could say if they wanted to mislead the neighbour or police about what actually happened. Finally, *cued counterfactuals* were measured by directly asking how things might have happened differently in each scenario. Throughout the study participants were allowed to look back at the scenarios as often as they wished. Specific wording for all questions asked is presented in Appendix 2B.

Outcomes of interest were coded as follows. Spontaneous counterfactuals were coded from the diary page text by two independent raters. The raters were not blind to hypothesis but coding for counterfactual thinking was done in the absence of knowledge about deception results. Counterfactuals were defined as thoughts about how events in the scenario could have been different (McEleney & Byrne, 2006). The scheme used to code counterfactual generations is included in Appendix 2C. Inter-rater reliability was high for both spontaneous ($r$ = .95) and cued counterfactuals ($r$ = .96). All discrepancies were resolved by discussion.

For spontaneous deception, participants coded their own statements. This was done in order to ensure that participants' deception was intentional and so that ambiguous statements (comments relating to the participants' own behaviour and traits) could be coded correctly. At the end of the study participants were instructed to look over their written statements in response to the two questions and code each sentence as either a lie or a truth according to the following definition of deception previously used by Serota, Levine, and Boster (2010):

> *'Most people think a lie occurs any time you intentionally try to mislead someone. Some lies are big while others are small; some are completely false statements and others are truths with a few essential details made up or left out. Some lies are obvious, and some are very subtle. Some lies are told for a good reason. Some lies are selfish, other lies protect others. We are interested in all these different types of lies.'*

Participants' classification of sentences into lies and truths was double-checked by one coder against the scenarios themselves. A lie was correctly classified as such only if it contradicted the events described in the scenarios. One participant mistakenly classed a truth as a lie; this was corrected by the coder. For each scenario, spontaneous and cued deception measures constituted the combined number of unique lies across the indirect and direct questions.

**Executive function.** Phonemic verbal fluency was tested with the *FAS test* (Strauss, Sherman, & Spreen, 2006). The test requires participants to generate as many words as possible beginning with the letters "F'', "A" and "S" within a one

minute time-frame for each letter. The score for each participant was calculated by adding together the correct words generated in response to each of the three letters.

Inhibitory control was assessed with the *Stroop Colour-Word test* (Golden & Freshwater, 2002; Stroop, 1935*)*. In the Colour task (C trial) participants had to read aloud, as fast as they could, words printed in incongruent colours. For the Colour word task (CW trial) participants were asked to name the colour of the font rather than the word itself. For each task, the time taken (seconds and milliseconds) to read all of the words was recorded. The ratio score (score from the CW trial divided by score from the C trial) was our outcome measure of inhibitory control (Sisco, Slonena, Okun, Bowers, & Price, 2016).

Cognitive flexibility was measured using the *Trail Making Test (TMT)* (Reitan, 1958). The *TMT* consists of two parts, in part A the participant has to connect circled numbers in a numerical sequence by drawing a line between them as rapidly as possible. In part B the participant is asked to draw lines to connect circled numbers and letters in an alternating numeric and alphabetic sequence as rapidly as possible. Using a hand-held stopwatch we recorded the completion time in seconds and milliseconds for each part. The TMT outcome score was calculated by dividing the score from part B by the score from part A (Golden, Osmond, Moses, & Berg, 1981).

Verbal working memory was tested using the Digit Span test from the Wechsler Memory Scale-Revised (WMS-R) (Wechsler, 1997). The test requires participants to repeat a sequence of numbers presented to them. For the Forward subpart participants need to repeat the sequence in the same order in which it was presented to them. For the Backwards subpart the sequence needs to be presented in

reverse. The Forward subpart includes a total of 16 sequences, the maximum amount of numbers in a sequence is 9. For the Backwards subpart there are a total of 14 sequences and the maximum amount of numbers in a sequence is 8. The total number of sequences correctly remembered for each subpart was added in order to obtain the score for this measure. Maximum score for the Digit Span measure was 30.

**Anxiety.** Anxiety was assessed using the Generalised Anxiety Disorder 7 (GAD-7) (Spitzer, Kroenke, Williams, & Löwe, 2006). GAD-7 is a self-report questionnaire which includes seven items which ask participants how often during the last two weeks they experienced anxiety related symptoms. Response options for each item ranges from 0 (not at all) to 3 (nearly every day). Total score was calculated by summing up all item scores.

## Results

For each outcome variable raw scores were transformed into z scores by subtracting the average from each raw score and dividing this by their standard deviation. Raw scores with a z score higher than 3.20 were treated as outlying cases and were removed from each of the outcome variables before any other analysis was performed (Tabachnick & Fidell, 2001). Based on this, a total of two participants were excluded, one from each of the outcome variables (less than 2.5% of data in each analysis)[2]. Assumptions of normality were tested using Shapiro-Wilk test and

[2] This procedure was used to identify outliers for every study in this thesis. From here on we will only mention outliers in instances where outliers were identified and removed.

visual inspection of Histograms and P-P plots[3]. Spearman's rho tests were used when assumptions of normality were violated.

**Association between scenarios**

Participants responded similarly to both scenarios. We found positive correlations between the 'moving town' and 'car incident' scenarios for measures of spontaneous and cued counterfactuals ($r$ = .31, $p$ = .005; $r$ = .33, $p$ = .003 respectively) and for measures of spontaneous and cued deception ($r$ = .26, $p$ = .023; $r$ = .49, $p$ < .001 respectively). Based on these results and the fact we were interested in averaging across different situational contexts we combined the equivalent counterfactual and deception variables.

**Association between counterfactual thinking and deception**

Table 2.1 shows the mean number of counterfactuals and lies generated across the two scenarios as well as the correlation between the variables of interest. As predicted, participants who spontaneously generated counterfactuals were more likely to generate spontaneous lies ($r$ = .24, $p$ = .034). Similarly, there was a correlation between the number of counterfactuals and lies when both were directly solicited ($r$ = .23, $p$ = .042).

[3] The assumption of normality was tested in this manner for all the analyses in this thesis. Non-parametric analyses were used unless data was normally distributed, or unless otherwise stated.

Table 2.1

*Means, standard deviations and correlations for counterfactuals and lies*

| | | | Correlations | | |
|---|---|---|---|---|---|
| | Range | *M* (*SD*) | 1 | 2 | 3 |
| 1. Total Spontaneous CFT | 0 to 6 | 1.73 (1.61) | | | |
| 2. Total Cued CFT | 2 to 13 | 7.10 (2.15) | .06 | | |
| 3. Total Spontaneous Lies | 0 to 9 | 2.73 (2.00) | **.24*** | .14 | |
| 4. Total Cued Lies | 2 to 15 | 8.31 (2.56) | .09 | **.23*** | .04 |

*Note:* The items in bold are our key comparisons. * $p < .05$

To ensure that this relationship is due to the availability of alternatives and not simply due to writing fluency, we also performed the above correlations partialling out the number of words that participants wrote in the diary task. These results were equivalent to the ones previously presented: ($r = .25$, $p = .032$) and ($r = .23$, $p =. 048$) for spontaneous and cued thoughts respectively, indicating a relationship independent of writing fluency.

As predicted, Table 2.1 also shows that there was no association between cued and spontaneous generations for either counterfactual statements or deceptive statements. Furthermore, we found no association between cued generations of counterfactual statements and spontaneous deception, or between spontaneous generations of counterfactual statements and cued generations of lies.

**Executive function, counterfactual thinking and deception**

Table 2.2 shows the association between executive function measures and counterfactual thinking and deception measures. In terms of counterfactual thinking,

the number of spontaneous generations was negatively associated with the number of words participants came up with on the FAS task. The more counterfactual thoughts participants generated in their diary pages, the fewer words they came up with in the response to the verbal fluency task.

Table 2.2

*Correlations between counterfactuals, lies and executive function*

| | Spontaneous CFT | Spontaneous Lies | Cued CFT | Cued Lies |
|---|---|---|---|---|
| FAS Test | -.28* | .09 | .20 | .28* |
| Stroop ratio score | -.04 | .10 | .12 | -.18 |
| TMT ratio score | -.01 | .13 | -.08 | .07 |
| Digit Span Total | -.02 | .06 | -.12 | -.09 |

* $p < .05$

We did not find any other associations between spontaneous counterfactuals and executive function. Cued counterfactuals also did not correlate with any such measures, although the correlation with the FAS task was approaching significance ($p = .079$). In terms of deception the tendency to come up with lies spontaneously was not associated with any of the measures assessing executive skills. The ability to come up with as many different lies as possible when asked to do so was however positively associated with the FAS. This implies that the more lies participants could think of the more words they would come up with on the FAS.

For spontaneous and cued generations in turn a single partial correlation which included scores on all executive functions measures at once was performed in order to understand whether these functions would impact the association between

counterfactual thinking and deception. The association between spontaneous generations of counterfactuals and lies remained positive and significant ($r$ = .33, $p$ = .005) whilst the association between cued generations of counterfactual and lies remained positive but was not statistically significant ($r$ = .18, $p$ = .126). A series of individual partial correlations which included each executive function score separately revealed that the association between cued generations of counterfactuals and deception wasonly influenced by Digit score and FAS score. Partialling out the influence of performance on Stroop and TMT tasks did not impact the significant association between cued generations.

**Anxiety, counterfactual thinking and deception**

Only 65 participants filled in the GAD-7, therefore the following results only relate to this sample. In this subsample of participants, the correlation between spontaneous generation of counterfactuals and lies was positive and statistically significant ($r$ =-.33, $p$ = .008), however the relationship between cued generations was only marginally significant ($r$ = .24, $p$ = .056). Overall, scores on the GAD-7 did not correlate with either spontaneous or cued counterfactual thoughts ($r$ =-.18, $p$ = .159; $r$ = .15, $p$ = .246, respectively). Results were similar for deception, anxiety was associated with neither spontaneous lies ($r$ = .05, $p$ = .705) nor cued lies ($r$ = .19, $p$ = .141).

In order to assess the impact of anxiety on the relationship between counterfactuals and deception we performed partial correlations between the generations of counterfactual and lies with anxiety as a covariate. Results suggest that the association between the spontaneous generations of counterfactuals and lies was not influenced by anxiety ($r$ = .34, $p$ = .006), the association between cued

generations only approached significance after controlling for anxiety ($r$ = .22, $p$ = .083).

## Discussion Study 1

As hypothesised, participants who came up with more counterfactuals were also more likely to come up with more lies. This was the case for both spontaneous and cued generations. These findings provide new insight into our understanding of the cognitive factors that are associated with the construction of deceitful statements.

The generation of counterfactual thoughts and lies were not associated in a similar manner with executive function skills. For counterfactual thinking these results stand in contrast with previous findings which suggest that aspects of counterfactual generations are associated with performance on tasks assessing inhibitory control and switching (Beck et al., 2009; Guajardo et al., 2009; Walsh et al., under review). Surprisingly, there was a negative relationship between phonemic fluency and spontaneous counterfactual, a finding which does not align with previous research which has emphasised a positive relationship between these skills (Solca et al., 2015).

In terms of deception, the cued generation of lies was associated with measures of phonemic fluency and this supports past research (Gombos, 2006). However it is surprising that spontaneous deception was not associated with any executive function measures, given that this type of deception has been found to rely heavily on these skills (Ding, Gao, Fu, & Lee, 2013).

The relationship between cued generations of counterfactuals and lies was affected by executive function, but the relationship between spontaneous generations

was not. The cued tasks relied more on verbal fluency because participants had to generate multiple alternatives. By contrast, the spontaneous generation task, especially in relation to deception, only required participants to generate one type of response. In addition, during the spontaneous tasks participants could look back over the scenarios and this could have diminished the need to keep information in mind, and thus reliance on working memory.

Overall, it could also be that that responses to scenario-based measures do not rely as much on executive skills, as responses were not time restricted and did not require interacting with others, an element which characterises real life situations (Buller & Burgoon, 1996). Whilst this is plausible, it has to be highlighted that previous studies which found a relationship between counterfactual thinking and executive function skills also used scenario based measures.

Another factor that should be taken into account is that there was little variation in our participants' executive function scores. Most of the studies which found a relationship between executive functions and counterfactual thinking or deception involved adults with an executive function impairment or young children (Abe et al., 2009; Beck et al., 2009; McNamara et al., 2003). Future work could investigate our hypothesis by assessing adults with atypical executive function. We will return to this in Chapter 5 where we will investigate the relationship between counterfactual thinking and deception in a population known to be impaired in executive function.

We also did not find any association between the variables of interest and anxiety. Once more, this was in contrast with previous research which found that anxiety is positively linked to counterfactual thoughts and deception (Prokopcáková

& Ruiselová, 2008; Van Der Zee et al., 2016). One potential explanation is that anxiety was assessed in relation to events in participants' own lives rather than in relation to the events described in the scenarios. Thus, any relationship between anxiety and our variables of interest might have been severed by this disconnect in the focus of the measurements used. In future, it would be useful to measure the role of emotions that derive from the same manipulation which also gives rise to counterfactuals and deception. We will seek to investigate this further in Chapter 4.

Whilst the underlying role of executive functions and emotions such as anxiety need to be investigated further, this is the first study to establish a direct link between counterfactual thinking and deception. As this is a correlational study, at this stage it is unclear what the underlying mechanism for this relationship is. It is reasonable to assume that counterfactual thoughts themselves enable participants to formulate a lie. This interpretation is congruent with that of Shalvi et al. (2011) who concluded that counterfactual thoughts offer people a self-justification which allows them to deceive whilst feeling honest. An argument that goes beyond that of self-justification is that counterfactual thoughts provide participants with the materials from which lies could be constructed. Equally, it could also be assumed that the need to deceive brings forward the need to think about what else could have happened, thus eliciting the construction of counterfactual thoughts. Such an interpretation would be in line with evidence which suggests that counterfactuals are activated by situational need (Roese & Epstude, 2017).

However, in line with Shalvi et al. (2011) interpretation we would like to further investigate our first assumption. Namely, we will aim to investigate whether the construction of counterfactuals can influence subsequent deception. The next two

studies aim to assess this experimentally by manipulating the availability of counterfactual thoughts and assessing whether this impacts subsequent deception.

## Study 2

Results from our first study show that individuals with a tendency to think counterfactually tend to generate more lies. In the current study we plan to examine the link between counterfactuals and deception by testing whether changing the availability of counterfactual alternatives also influences the tendency to lie. We know that our intrinsic honesty is compromised by characteristics of the environment and of the situation we find ourselves in. In environments where the prevalence of rule violation is high people tend to lie more (Gächter & Schulz, 2016), and dishonesty is enhanced if the degree to which the truth needs to be altered is small (Hilbig & Hessler, 2013).

The characteristics of a particular situation also affect the generation of counterfactual thoughts. Counterfactual thoughts are more likely following unusual and near-miss situations because in these situations the antecedents have alternatives that have been previously experienced and are therefore more easily imagined. Following an unusual situation, individuals are more likely to recapitulate the normal state of events by thinking of a what would have usually happened (i.e. if involved in an accident on a new route home one might think ‘If only I had taken my usual route, the accident wouldn’t have happened’) (Kahneman & Miller, 1986). Following a near-miss situations an individual is more likely to draw a comparison between the actual and the missed outcome and imagine the things they could have changed in order to obtain the desired outcome (e.g. after missing the train by 5 minutes one might think: ‘If I had run faster I could have caught the train’). In other words,

unusual and near-miss events are more cognitively mutable and thus counterfactual thought are more frequent in these contexts. If generating counterfactuals is an important component in deception, we would expect to find more lies in these situations where counterfactuals are readily available. In this study we investigate this specific hypothesis using two versions for each scenario used in Study 1, one with a highly available counterfactual alternative and one without. We expected that people will generate more lies when the events are presented as either unusual or as if they almost didn't happen, due to the increased availability of counterfactuals.

## Method

### Participants

A total of 50 participants (35 female and 15 male) were recruited from students at Plymouth University and given course credit for their participation. Participants were aged between 18 and 39 years (*M* = 20.26, *SD* = 3.39). The sample size was limited to the available number of Plymouth University students that had already taken part in Study 1[4].

### Materials and Procedure

The scenarios used in study 1 were modified slightly for the purposes of this study. We manipulated the availability of counterfactuals related to participants' motives for deception. In the first scenario participants tended to lie about going to the cinema because this was seen as a bad excuse for not attending the party, we

[4] We tried to limit participant overlap as much as possible between the studies in this thesis. For studies conducted in the laboratory this was guaranteed by limiting the advertisement of the study to participants that had not taken part in any previous study. Studies conducted online did not require participant's names and thus the same strategy could not be used, however when using the same recruitment website for different studies the study was advertised under the same name to limit potential participant overlap.

therefore manipulated the elements surrounding this decision. In the second scenario participants lied to the police in order to avoid being found guilty of the incident and thus we manipulated the counterfactual availability of alternatives relating the incident itself.

Participants were randomly allocated to two conditions: either a low counterfactual condition or a high counterfactual condition. In the low counterfactual condition events were manipulated to appear usual and temporally far (i.e. the decision to go to the cinema is something they do every week; the car that they end up scratching had been parked there for a long time). In the high counterfactual condition, the events that lead to the deception were manipulated to appear unordinary and temporally close (i.e. the decision to go to the cinema by themselves is something they never usually do; the car they end up scraping had only parked there only minutes before). The scenarios were virtually the same as the ones in Study 1 (Appendix 2A) apart from the fact that they included the changes stated above.

The order of the scenarios was counterbalanced, but counterfactual and deception tasks were always presented in the same order. First participants read the modified scenarios and wrote a dairy page, in a similar way to Study 1. After writing each of the diary pages, participants responded to two questions aimed to assess their propensity to deceive. To assess cued generations of counterfactuals participants were asked to provide as many thoughts about how this could have happened differently. For cued deception participants were asked to come up with as many lies as they could think of (these questions were identical to the ones in Study 1 and are presented in Appendix 2B). Next, participants completed a series of self-report

questionnaires in order to assess baseline tendency to engage in counterfactual thinking and deception. The Schwartz Regret Scale (Schwartz et al., 2002) and the self-report measure used in Beldarrain et al. (2005) study were used to assess aspects of counterfactual thinking. Participants were also asked to report the number of lies they told in the last 24 hours to either family, friends, acquaintances. We recorded the total number of lies reported across all of the different situations Serota et al. (2010). Lastly participants coded their own responses as truthful or deceitful according to a definition of deception by (Serota et al., 2010). Counterfactuals and lies were coded just as in Study 1.

## Results

Outlying cases with z > 3.20 were removed from each of the outcome variables before any other analysis was performed (Tabachnick & Fidell, 2001). Based on this, for the spontaneous counterfactuals variable in the car crash scenario, one participant was removed from the low counterfactual condition (2% of data in one of the analysis).

### Baseline counterfactuals and deceptive tendencies

The two groups did not differ in their baseline tendencies to engage in counterfactual thinking as measured with the Schwartz scale ($M$ = 22.60 in the Low CFT condition vs $M$ = 21.88 in the High CFT condition, $U$ = 284.50, $p$ =.756) or as measured with the Gomez-Belderrain scale ($M$ = 17.72 in the Low CFT condition vs $M$ = 16.33 in the High CFT condition, $U$ = 258.50, $p$ = .406). Similarly, the two groups also did not differ in their in their natural propensity to deceive as indexed by the number of lies they had told in previous 24 hours ($M$ = 2.64 in the Low CFT group vs $M$ = 1.92 in the High CFT group, $U$ = 229.00, $p$ = .148).

**Counterfactual manipulation check**

In a similar manner to Study 1 we combined the counterfactual scores for both scenarios. An analysis of differences between scenarios within each condition revealed no significant difference between the way in which participants responded to each scenario (all ps > .066).

Descriptive statistics for cued and spontaneous generations in each of the two conditions are shown in Table 2.3. Participants in the two conditions did not differ significantly in regards to the number of spontaneous counterfactuals they generated $U = 277.0$, $p =. 475$. Similarly participants did not differ in terms of cued generations, $U = 236.5$, $p =. 287$.

Table 2.3

*Means and standard deviations for spontaneous and cued counterfactual responses in each experimental condition*

| | Condition | |
|---|---|---|
| | High CFT | Low CFT |
| | *M (SD)* | *M (SD)* |
| Spontaneous counterfactuals | 1.36 (1.55) | 1.52 (1.39) |
| Cued counterfactuals | 5.30 (2.14) | 5.96 (1.88) |

**Deceptive responses**

As in Study 1 we added together the deception scores form each scenario. An analysis of differences between scenarios within each condition revealed no

significant difference between the way in which participants responded to each scenario (all p's > .733).

The mean number of spontaneous and cued lies for each condition is shown in Table 2.4. Overall participants in the high counterfactual condition did not come up with more lies in comparison to the participants in the low counterfactual condition $U = 288.00$, $p = .630$. In terms of cued generations, participants in the High CFT group came up with significantly less lies than those in the Low CFT group ($U = 188.00$, $p = .037$).

Table 2.4

*Mean and standard deviations for spontaneous and cued lies across experimental conditions*

| | Condition | |
|---|---|---|
| | High CFT | Low CFT |
| | *M (SD)* | *M (SD)* |
| Spontaneous deceptive generations | 4.44 (1.86) | 4.68 (1.86) |
| Cued deceptive generations | 6.08 (2.33) | 7.45 (2.28) |

## Discussion Study 2

Participants in the two experimental groups did not generate a significantly different number of counterfactual thoughts, thus our experimental manipulation did not work as intended. There was a trend in the opposite direction to the one predicted, namely, participants in the low counterfactual condition tended to produce

more counterfactual than those in the high counterfactual condition, both in terms of cued and spontaneous responses. The trend was similar for deceptive responses, and for the cued generations participants in the Low CFT group produced significantly more lies than participants in the High CFT group. However given that our manipulation did not work as intended we cannot draw a conclusion about the influence of counterfactual saliency on subsequent deceptive communication.

Considerations relating to the way in which the scenarios were modified for this study could explain why the experimental manipulation was unsuccessful. The scenarios were only minimally modified, that is we only modified one element in relation to the event of interest, whilst keeping all other elements equivalent. The scenarios also included other events, which were not manipulated and which could have led participants to counterfactual thoughts. For example, in the moving town scenario, apart from the decision to go to the cinema instead of the neighbour's party, participants could also undo the decision to move town or the decision to buy a new stereo. Similarly, in the car incident scenario, the participants could also undo their decision to drive down a new route or the decision to answer the phone whilst driving. Thus, regardless of our experimental manipulation counterfactuals might have been readily available in both conditions. It could also be that the explicit assessment of counterfactual thoughts through the diary page prompted participants to come up with counterfactual thoughts independent of the experimental manipulation.

To improve the methodology participants could be presented with shorter scenarios that only included one event which can then be directly manipulated in

terms of the counterfactual thoughts it can elicit. The study which follows was designed accordingly.

## Study 3

In Study 2 we attempted to investigate whether situations known to bring about counterfactual thoughts would lead to an increase in deception. Our experimental manipulation was not successful and thus a conclusion about whether counterfactuals are causally linked to deception could not be drawn. In the current study we aimed to ask the same question as in Study 2, but this time we attempted to modify the experimental design in order to address the previous limitation. Specifically, we used shorter versions of the original scenarios which included a single event leading to a negative outcome. The event in each scenario was manipulated in order to elicit readily available counterfactuals. As before, we expected people to generate more lies when the events were presented as unusual or as if they almost didn't happen, as these are situations that make counterfactuals readily available.

## Method

### Participants

One hundred and twenty-two participants (53 female and 69 male) were recruited online. Participants were not recompensed for their participation. They ranged in age from 18 to 50 years ($M$ = 24.38, $SD$ = 6.65). The study had a between subjects design and participants were randomly assigned to the low counterfactual condition (n = 60) or to the high counterfactual condition (n = 62). An a priori power analysis showed that 59 participants per condition should have 80% power to detect

an effect size Cohen's d of 0.50. We collected data until the close of day on which the actual sample size reached or exceeded 59 in each condition.

**Materials and Procedure**

We constructed high counterfactual and low counterfactual versions of the scenarios used in study 1. Participants received either both high counterfactual versions or both low counterfactual versions in a counterbalanced order. As before, participants were asked to imagine themselves in the scenario. To make the differences more salient, each scenario contained only one explicit decision and the two versions of the scenarios differed only in the mutability of that decision. For the 'moving town' scenario, participants were solely presented with the decision to go the cinema rather than the neighbour's party. As in Study 2, in the low counterfactual version, the decision to go to the cinema was portrayed as routine (i.e. something done on a regular basis), whilst in the high counterfactual version, the decision to go to the cinema was described as exceptional (i.e. something never done).

We removed the diary page task from this study in order to limit participants' tendency to think of counterfactuals beyond the events presented in the scenario. After reading the 'moving town' scenario participants had to respond to questions testing for spontaneous deception. As opposed to Study 1 and 2 we chose to include only the direct deception questions. This was done in order to avoid repetition. In particular, for the 'moving town' scenario participants were asked to write down anything they would say to the neighbour if asked why they failed to attend the

party. To measure cued deception, they were then asked to list all the lies they could tell the neighbour in that instance.

For the ‘car incident’ scenario, as in Study 2, the low counterfactual version mentioned that the car that was hit had been parked there for a long time, whilst in the high counterfactual version, the car had only parked there minutes before. After reading the scenario participants had to write anything they would say to the police when they were asked whether they knew anything about the car incident. Following this question participants were also directly asked to list as many lies as they could think of in order to mislead the police.

Finally, participants completed self-report measures of counterfactual thinking and deception. The Schwartz Regret Scale (Schwartz et al., 2002) and the self-report measure used in Beldarrain et al. (2005) study were used to assess aspects of counterfactual thinking and the self-report measure used in the Serota et al. (2010) study to assess the number of lies participants told in the past 24 hour.

**Coding**

Two independent raters which were not blind to hypothesis but blind to experimental condition coded the number of deceptive responses given in response to each of the two scenarios. Deceptive responses were coded based on the same definition used in Study 1, additionally the coding scheme used in appendix 2D was also used. Each lie was given a score of 1 and inter-rater reliability was high for both spontaneous ($r$ = .94) and cued deception ($r$ = .96).

## Results

Outlying cases with z > 3.20 were removed from each of the outcome variables before any other analysis was performed (Tabachnick & Fidell, 2001). Based on this, for the spontaneous deception outcome variable, two participants were removed from the low counterfactual condition and three participants were excluded from the high counterfactual condition, for the cued generations, two participants were removed from the low counterfactual condition (4% of data in each analysis).

### Baseline counterfactual and deception tendencies

The two groups did not differ in their baseline tendencies to engage in counterfactual thinking as measured with the Schwartz scale ($M$ = 21.97 in the Low CFT condition vs $M$ = 21.35 in the High CFT condition, $U$ = 1776.0, $p$ = .560) or as measured with the Gomez-Belderrain scale ($M$ = 17.54 in the Low CFT condition vs $M$ = 16.66 in the High CFT condition, $U$ = 1742.0, $p$ = .450). Similarly, the two groups also did not differ in their propensity to deceive as indexed by the number of lies they had told in previous 24 hours ($M$ = 4.75 in the Low CFT group vs $M$ = 3.33 in the High CFT group, $U$ = 1625.0, $p$ = .355).

### Deception Scores

As in our previous studies scores for each scenario were added in order to analyse the overall deception scores. A comparison between scenarios in each of the experimental condition revealed no significant differences (all p's > .389). Mean deception scores for each condition are presented in Table 2.5. A comparison between the two conditions revealed that participants in the High CFT condition spontaneously produced more lies ($U$ =1347.0, $p$ = .025, $r$ = .21). The two groups

did not differ however in cued generation of deceptive responses ($U$ = 1658.0, $p$ = .366).

Table 2.5

*Means and standard deviations for spontaneous and cued lies across experimental condition*

| | Condition | |
|---|---|---|
| | High CFT | Low CFT |
| | *M (SD)* | *M (SD)* |
| Spontaneous deceptive generations | 1.93 (1.19) | 1.47 (.92) |
| Cued deceptive generations | 3.90 (2.13) | 4.17 (2.05) |

## Discussion Study 3

Results show that people tend to spontaneously generate more lies when counterfactual alternatives are readily available. So far the studies in this chapter suggest that counterfactual thinking and deception are positively associated and that counterfactuals lead to deception. These findings are consistent with results from Shalvi et al. (2011) and further suggest that simply imagining a counterfactual alternative is sufficient in order to influence deceptive communication.

A different line of evidence for the relationship between deception and counterfactual thinking has come from Miller et al. (2005) who have shown that spontaneous counterfactual thoughts are positively associated with inferences of deception. In order to support the link between counterfactual thought and deception we therefore also wanted to assess whether inferences about a person's

counterfactual thought would also lead to inferences about deception. The two studies which follow aim to assess the relationship between counterfactual thinking and deception from this different vantage point.

## Study 4

Counterfactuals can influence the way in which we judge others. The contrast between what actually happened and what could have happened can give rise to conclusions about what caused the event and can in turn influence the ascriptions we make in regard to other people's intentions, traits or emotions (Gavanski & Wells, 1989; Kahneman & Tversky, 1982b; Knobe, 2010). For example, imagine hearing about Ben, who was attacked while out running. If he took a new route that day then people may readily imagine a scenario in which he took his usual route and wasn't attacked. Exceptional events such as this are often the focus of counterfactual thoughts and as a result are often blamed for unwanted outcomes and can amplify emotional reactions (Kahneman & Miller, 1986). The salience of what could have happened which is brought through by unusual events and near-misses has the ability to bias our judgements.

It then follows that, if as suggested by our previous studies, lies are more likely in situations where counterfactuals are more readily available, then such circumstances should also influence our perceptions of *other's* intentions to lie. Therefore, the current study tested whether people will also be judged as more likely to lie in situations where counterfactual alternatives are more salient. Miller et al. (2005) have already shown that we tend to judge others as more dishonest if we ourselves imagine how they would have done so under different circumstances. However, in the current study we were interested to see whether there is a

relationship between the protagonist's own counterfactual thoughts and the likelihood that they will be judged as deceptive.

The study was modelled on the Counterfactual Inference Test (CIT, Hooker et al., 2000), a multiple choice test used to measure people's tendency to use available counterfactuals in making affective judgements. In this study, we examined whether available counterfactuals are also used to make judgements about the likelihood that someone will lie.

## Method

### Participants

A total of 102 participants (43 female and 59 male) were recruited and performed the study online. They ranged in age from 18 to 43 ($M$ = 23.71, $SD$ = 5.49). Participants were not rewarded for their participation.

### Materials and Procedure

We generated six scenarios and similarly to our previous studies, we used unusual events and near-misses to elicit counterfactual thinking. Each scenario described similar events experienced by two actors which led to the same negative outcome. For half of the scenarios, one action was exceptional and the other routine, for example:

> *John crashes his dad's car while driving on his usual way home. Bob crashes his dad's car whilst trying a new way home.*

The other three scenarios also contained two similar events but one was a near miss whereas the other was not, for example:

*Adele and Rachel are taking part in an experiment. They are given a die and told that if they roll a six three times they will win ten pounds. The experimenter cannot see the actual die rolls. Adele rolls two sixes, Rachel rolls one six.*

The scenarios included a third party which could subsequently be the target of an act of deception. In the examples above the third parties are the dad and the experimenter respectively. A full set of scenarios are presented in Appendix 2E. The scenarios were presented twice in two separate blocks. The counterfactual questions were administered in one block and the deception questions were administered in a separate block. The two blocks were separated by a set of questionnaires which were not of interest to this study. The presentation order of the blocks was randomised as were the order of items within the blocks and the order of response options.

To assess whether participants attributed more counterfactual thoughts to the protagonist that experienced the unusual event or the near-miss, all items asked the following question:

*Who spends more time thinking about how things could have been different?*

The questions which assessed participants' inferences about the protagonists' honesty all had the following structure:

*Who is more likely to lie to x about y?*

where x denotes the target of deception in the scenario, and y was the outcome. For the first example above, the question was:

*Who is more likely to lie to their dad about crashing the car?*

All questions had three response options which were identical for each section. One option was the protagonist associated with the target counterfactual event, another one referred to the alternative protagonist in the scenario and a third final option was 'Same/can't tell'. For the above item the target response option was 'Bob'.

**Pilot Study**

To check if the scenarios elicited counterfactual inferences, we pilot-tested the counterfactual questions on a sample of 28 participants. For each item, the majority of participants (68% to 92%) selected the response option that referred to the target counterfactual response more often than the low counterfactual alternative. A chi-square analysis for each item revealed that this pattern was significant ($\chi^2$ (2) = 14.48, 21.44, 25.04, 38.96, 29.12, 13.76; all *p*'s <.001). These preliminary results suggest that all items elicited counterfactual inferences in the majority of participants.

## Results

Results were checked for outliers, but none were identified therefore all data were included in the analysis. Data was normally distributed therefore we used parametric analyses.

Table 2.6 shows the percentage of participants choosing each response option for the six items. We were specifically interested in the endorsement rates for the target counterfactual options versus the non-target alternative response. Participants

were significantly more likely to choose the target option than the non-target option in response to both the counterfactual (67% vs.7%; $\chi^2(1)= 293.13$, $p < .001$) and deception questions (51% vs.13%; $\chi^2(1)= 135.64$, $p < .001$). This pattern occurred for the counterfactual question in all six scenarios and for the deception question in all scenarios except scenario 6.

The results also showed a positive correlation between counterfactual and deception judgements ($r = .39$, $p < .001$). When participants endorsed the target option for the counterfactual question, they tended to do so for the deception question also.

## Discussion Study 4

The results of this study are consistent with the results of our previous studies and suggest that people use counterfactual inferences in order to make judgements relating to dishonesty. Overall, participants chose the counterfactual target option when asked about the protagonists' likelihood to lie. This pattern occurred for the counterfactual question in all six scenarios and for the deception question in all scenarios except scenario 6. One explanation for the exception in item 6 may be that the causal link between the event which could be undone (i.e. getting off at the unusual vs usual bus stop) and the outcome (forgetting the package on the bus) was weaker in comparison to the other items. This highlights the fact that additional factors may be needed for individuals to use counterfactuals to lie.

Table 2.6.

*Percentages of participants endorsing each response option each of the six scenarios.*

| Scenario | Questions | CFT response | Alternative response | Same/ can't tell |
|---|---|---|---|---|
| 1 | • Counterfactual | 60% | 8% | 32% |
| | • Deception | 50% | 5% | 45% |
| 2 | • Counterfactual | 76% | 4% | 20% |
| | • Deception | 68% | 9% | 23% |
| 3 | • Counterfactual | 80% | 10% | 10% |
| | • Deception | 76% | 11% | 19% |
| 4 | • Counterfactual | 71% | 5% | 24% |
| | • Deception | 36% | 17% | 47% |
| 5 | • Counterfactual | 71% | 6% | 23% |
| | • Deception | 60% | 11% | 29% |
| 6 | • Counterfactual | 44% | 10% | 46% |
| | • Deception | 21% | 27% | 52% |
| Total | • Counterfactual | 67% | 7% | 26% |
| | • Deception | 51% | 13% | 36% |

Although highly likely, it is unclear whether in fact participants used the counterfactual thought relating to this protagonist in order to make the assumption of dishonesty. Participants could have used other rationales, such as inferences about the protagonist's character. For example, for item 2, participants could have inferred that Sarah is more likely to deceive because she is careless in her decision to unexpectedly stop for coffee, which might then further imply that she is overall of a questionable character. In such an instance participants' choices would not have been directly related to the protagonist and not the counterfactual nature of the situation. The study which follows seeks to clarify by investigating the reasons underlying people's choices in this study.

## Study 5

Our previous study showed that we ascribe dishonest behaviour to people in situations where counterfactuals are readily available. In the current study we attempted to replicate the results of study 4 whilst also examining the reasons why people are more likely to choose the counterfactual target response. It was hoped that these findings would provide us with insights as to the mechanism linking counterfactuals and deceptive inferences.

### Method

#### Participants

A total of 89 participants (59 female and 30 male) were recruited and performed the study online. They ranged in age from 18 to 54 ($M$ = 23.52, $SD$ = 6.02). Participation in this study was not rewarded monetarily or otherwise.

## Materials and Procedure

We used the same six items as in Study 4, however this time participants were only administered the deception subpart of the questionnaire. In addition to our previous study the current questionnaire also included a question which asked participants to state the reasons behind their choice. The order in which the items were presented was randomised as was the order of the response options for each item.

The reasons participants gave for their choices were classified as either situation-specific, lie-specific, or character-specific. Reasons which related to the counterfactual nature of the situation were labelled as situation-specific. Explanations which related to the characteristics of the lie to be told were labelled as lie-specific, and those which referred to characteristics of the protagonist themselves were classed as character-specific. Each explanation was assigned to only one category.

# Results

## Endorsement of target responses versus non-target responses

Participants were significantly more likely to choose the target option than the non-target option in response to the deception questions (39% vs.13%; $\chi^2(1) = 71.77$, $p < .001$). Again, this pattern occurred for all six scenarios except scenario 6, where participants picked the non-target option more than the target option (21% vs 10%).

**Reasons for endorsing target responses**

When looking at the reasons participants provided for choosing the counterfactual target option the majority of responses (62%) gave reasons which were specific to the counterfactual nature of the situation. In other words, the majority of participants reasoned that the heightened counterfactual aspect of the situation would elicit the protagonist to deceive. For example, for scenario 1, some participants reasoned that Adele would be more likely to lie because: ’She was closer to the goal of three sixes than Rachel.’ For scenario 2, some participants judged that Sarah, the protagonist placed in an unusual situation would be the one more likely to lie because: ‘It was out of her routine. Alison may feel like coming home always includes coffee, so it may not stand out to her as something she did wrong’.

Some participants also explained that they selected the target option because a lie in such a situation would feel smaller and thus easier to tell. In total 20% of reasons were based on the characteristic of the lie that would be told. For example, for scenario 3, a participant stated that Jenny would be more likely to lie ‘Because the lie will be smaller and easier to do’, similarly for scenario 4 a participant stated that Ed would be more likely to lie because: ‘Ed feels like his lie is more believable’. Also, responses included the judgement that ‘Rachel does not think the lie is too big’ for scenario 1.

Only 3% of participants referred to the characteristic of the protagonist as a reason for deception. For example, a participant chose Jenny in response to scenario 3 because ‘she likely has enough relevant skill to lie’. The remaining 15% of participants did not provide a reason for their choice.

Interestingly, out of the reasons which related directly to the counterfactual nature of the situation, 16% directly referenced that the counterfactually derived emotions experienced by the protagonist would affect their decision to lie. For example, in response to scenario 2, a participant chose the counterfactual target response Sarah because 'She feels guilty because she feels her change in pattern created the problem'. An example of a similar interesting response is the following one in relation to scenario 5: 'Bob might feel guilty for choosing the risky route, and thus want to have an excuse or to hide that fact'. Also, for scenario 4 some reasoned that Ed would be more likely to lie because 'He would feel bad about missing it by so little time'. Emotion related responses were most frequent in response to Scenario 2.

## Discussion Study 5

Overall, the results of this study replicate the results of Study 4. The overall attribution of deception to the counterfactual targets was lower than in Study 4 (39% vs 51%, respectively). A potential explanation is that questioning participants about their rationale lead to a deeper consideration of the scenarios and thus reduced the bias. The study however enhances findings from Study 4 by showing that people do indeed use counterfactual information in order to make judgements relating to dishonesty.

Similar to results in Study 4, we found that for Scenario 6 participants attributed increased deception tendency to the protagonist experiencing low counterfactual availability. Explanation provided by participants help interpret this finding. Most participants explained that Steve would be more likely to lie because Aaron (the protagonist who got off at a different stop to the one they usually get off

at) could use the change in pattern as an excuse for the negative outcome. This was also seen as a reason for why Aaron would be less responsible for the outcome. For example one participant stated ‘More things are likely to wrong when doing something new so Aaron wouldn’t get all the blame’. This explanation highlighted fact that counterfactuals may elicit truth-telling if focused on elements which reduce blame.

Interestingly, some participants mentioned that emotion derived through the counterfactual nature of the situation could lead to an increase in deception. These findings are congruent with the evidence about the driving force of emotion in deceptive communication. For example, Depalma, Madey, and Bornschein (1995) found that guilt which follows a misdeed can increase deception. This might indicate that the mechanisms underlying the relationship between deception and counterfactual thinking is partly motivational. We will attempt to further investigate whether emotions can influence the link between counterfactuals and deception in Chapter 4.

Additionally, we found that people consider that lies which follow a situation where counterfactuals are more available as easier to tell and more likely to be credible. These findings suggest that people use closeness to judge the believability of a lie. These findings align with the interpretation by Shalvi et al. (2011) according to which counterfactuals enable self-justifications as to why lying is ethically acceptable in some situations. These results are also congruent with evidence which suggests that a reduced difference between the truth and lies has the potential to increase deception (Hilbig & Hessler, 2013).

## General Discussion

The current studies in the current chapter are the first to directly investigate the influence of counterfactual thinking on deceptive communication. Study 1 showed that individuals with a tendency to think counterfactually are more likely to generate lies. Our results also showed that when counterfactual alternatives are highly available, people are more likely to generate lies (Study 3), and to judge that others will lie (Study 4 and 5). Overall, the findings from all these studies are the first to directly emphasize the positive link between counterfactual thinking and deception. Additionally, these findings are also the first to demonstrate that merely imagining an alternative scenario is sufficient to encourage deception.

Counterfactual thoughts represent an account of what might have been and therefore provide a salient alternative which could be used to generate a lie. But, in our studies participants did not solely use the content of their counterfactual thoughts in generating deceptive responses, therefore the effect of counterfactual thoughts on deceptive communication may be in part through a content-neutral pathway (Epstude & Roese, 2008). Content neutral effects emerge when counterfactual thoughts ignite attentional and cognitive processes that further influence behaviour independent of the specific information and meaning contained in the original thought. We propose that counterfactuals play a role in deception by making useful information available to the deceiver. Namely, we want to argue that thinking about how a specific event could have been different increases the general availability of alternatives to past events which in turn aids the generation of lies.

In Study 1, our measures of deception and counterfactual thinking were not associated with measures of executive functions. This could be because the student

population we tested varied little in these skills. However, this should be further investigated given that we did find the relationship between cued generation of alternative possibilities and lies to be affected by certain executive functions. This is also important given that other studies emphasized the strong role of such skills in the generation of counterfactual thinking and deception (Beck et al., 2009; Gombos, 2006; Guajardo, McNally, & Wright, 2016; Walsh et al., under review). Chapter 5 will seek to address this further.

Beyond the cognitive links, counterfactual thinking may help to justify deception because events that could have been true may seem more plausible. For example, Shalvi et al. (2011) found that lies about the outcome of a die roll tended to be regarded as less unethical when the lie matched an alternative observed roll. Lies that only differ a little from the truth also feel more honest and are judged as more justifiable by others (Lelieveld, Shalvi, & Crone, 2016). Similarly, Schweitzer and Hsee (2002) found that car sellers were more likely to give an exaggerated estimate of the unknown mileage of a car when the estimate fell within a range of possible values. Therefore those who think that things happened nearly as they desired might lie about it more because they believe this is admissible to themselves as well as others. Results in Study 5 which show that lies following a counterfactually heightened situation are judged as more believable support this reasoning.

A motivational mechanism that also needs to be considered is affect. In Study 1 contrary to previous research anxiety was not associated with either deception or counterfactual thinking scores. However in Study 5 some participants did base their inferences of deception on emotions which were derived from the counterfactual nature of a situation. Therefore, it could be that counterfactually derived affect also

plays a role in the relationship between counterfactuals and deception. Counterfactual thoughts are known to elicit a wide range of affective reactions, such as guilt and shame (Niedenthal, Tangney, & Gavanski, 1994) and some studies suggest that increased post-transgressional guilt can lead to more lying (Depalma et al., 1995). We will return to examine this potential mechanism in Chapter 4.

Independent of the mechanisms that characterise this relationship, our results have implications for the fields of both counterfactual thinking and deception. Firstly, they provide insight into the possible functions of counterfactual thinking. Its primary function is thought to be in helping people to learn from past mistakes and develop intentions for the future (Epstude & Roese, 2008). However, people often generate counterfactuals focussing on uncontrollable events which cannot be used to improve on the past (Ferrante, Girotto, Straga, & Walsh, 2013; Mercier et al., 2016) suggesting that counterfactual thinking may also serve other functions. Our findings suggest that one of those functions is to enable people to deceive. Given that lying can constitute an important phenomenon in interpersonal relationships, with some regarding it as a 'social lubricant' (Vrij, 2007), our results highlight the importance of counterfactual thinking for social functioning. The idea is consistent with findings that reasoning serves the social function to persuade others (Mercier & Sperber, 2011). Secondly, these findings also contribute to the field of deception by helping to identify the processes that underlie it (Gamer & Ambach, 2014). Our current findings suggest that counterfactual thinking is one of those mechanisms. These findings also contribute to the emerging discussion regarding the role of individual differences in lying behaviour. Understanding the elements that characterise a frequent liar is of great relevance to the field of human communication. Currently, most theories of deception specify the role of basic cognitive processes such

inhibitory control, working memory and response monitoring (Vendemia, Schillaci, Buzan, Green, & Meek, 2009; Walczyk, Roper, Seemann, & Humphrey, 2003) however studies such as the current one emphasise that in order to develop deception theories we need to also model the interplay between such general mechanisms and individual characteristics.

In conclusion, the studies in this chapter provide the first direct link between people's tendency to think about alternatives to the past and lying. As a result we have new insight into how individual differences in the tendency to generate counterfactuals may influence people's behaviour. Furthermore this research also identifies an additional cognitive process involved in deception. Overall, we have provided an initial step towards understanding the link between imagination and lying.

# CHAPTER 3

## Do counterfactuals influence individuals' real-life deception?

Our previous studies used hypothetical scenarios to show that people are more likely to generate lies and to judge that others will lie in situations which also tend to elicit counterfactual thoughts. Given that so far we have used scenario-based methodology, in this chapter we aim to extend our findings by using methodology which more closely resembles real-life situations.

Participants in our studies could easily and readily engage with the scenarios as evidenced by their inclusion of names and description of feelings, however even though participants were asked to vividly imagine themselves in certain situations, the hypothetical nature of the scenarios may not have replicated the multifaceted nature of real-life deception. Firstly, participants in our previous studies were not subjected to any time pressure when answering deception questions. In contrast, in most real life situations little time is allowed for such reflection and deception has been found to be dependent on the time allotted for a response. For example, Gunia, Wang, Huang, Wang, and Murnighan (2012) found that participants lie more when required to respond immediately rather than when asked to carefully contemplate whether to lie or not. Similarly, Shalvi, Eldar, and Bereby-Meyer (2012) found that people are more likely to lie by using self-serving justifications such as observed counterfactuals in situations where time is limited.

Secondly, the medium through which deception is communicated is also an important factor. In our previous studies, although the participants had to imagine a face-to-face conversation, their responses to the deceptive questions were given in

writing. Written communication has been shown to attenuate cognitive effort by eliminating the need to monitor visual and auditory cues, thus enabling participants to create more believable and effective messages (Burgoon, Stoner, Bonito, & Dunbar, 2003). In support, Zimbler and Feldman (2011) have shown increased deception in computer-mediated communication such as email and instant messaging rather than face-to-face interactions.

Lastly, our hypothetical scenarios might have minimised the incentive to deceive which would normally characterise most real-life deception situations. Our scenarios attempted to account for both high-stakes (lying to a police in order to avoid legal punishment) and low-stakes (lying to a neighbour in order to avoid social punishment) motivational factors. However, as these situations were hypothetical, participants' choice to lie or tell the truth was inconsequential. Evidence on the influence of increased motivation to lie on production success has been mixed. Some argued that increased motivation can impair the control of non-verbal cues, and as a consequence can render lies more transparent (DePaulo & Kirkendol, 1989). In particular, ten Brinke and Porter (2012) argue that a high motivation to deceive successfully is accompanied by powerful emotions, such as fear and remorse which need to be monitored and inhibited, making such deception more cognitively demanding. Conversely, others have suggested that increased motivation can in fact facilitate a deceiver's verbal performance (Bond & DePaulo, 2006). Such research argues that high motivation to deceive should make the rationale behind lying obvious and thus in such a situation one should rarely feel guilty or ashamed for deciding to lie. Either way, motivation is an important factor when lying and should be taken into account when designing paradigms which aim to assess this type of

communication. Overall these considerations highlight the reasons why assessing deception with hypothetical scenarios might be limiting.

Counterfactual thinking can also be influenced by the way in which information is presented to participants. Whilst most studies investigate how readers think about fictional situations, Girotto et al. (2007) have shown that participants' counterfactual thoughts can differ depending on whether participants read about a situation or actively participate in it. For example, those who read hypothetical scenarios tend to undo the protagonists' decisions, whereas those that are actually involved in similar scenarios tend to undo elements pertaining to the situation itself. Observing an event is also akin to acting it out in terms of the counterfactual thoughts which follow (Pighin, Byrne, Ferrante, Gonzalez, & Girotto, 2011) but different from reading about the event. These differences are explained by the fact that attention is engaged differently in the different roles, and thus different information becomes available to participants which in turn shapes the content of counterfactual thoughts. Given that in our past studies participants were readers, the question that remains is whether the information which would be available through experiencing an event would influence deception in a similar way.

Overall there are several factors important in both deception and counterfactual thinking which have not been taken into account in our previous studies. The current chapter will attempt to address these limitations and aim to assess whether the link between counterfactual thinking and deception persists when using experimental paradigms which more closely resemble real-life situations. In Study 6 we aim to investigate the association between counterfactual thinking and deception when lying involves face-to-face communication under time pressure. In

Study 7 we will aim to allow participants to experience the counterfactual nature of an event whilst also increasing their motivation to deceive a co-player. The current chapter will allow us to conclude whether people use counterfactuals thoughts to deceive when given a real-life opportunity to do so.

## Study 6

The current study will utilise a deception testing paradigm with increased ecological validity. As mentioned previously the medium of communication and the time allotted for the generation of deceptive messages needs to be considered when studying lying**.** The Deception Interaction Task (DeceIT) paradigm developed by Wright, Berry, and Bird (2012) takes into account these factors by assessing lying within an interactive group-based competitive game. The paradigm requires individuals to verbally lie in front of others and is advantageous because it allows for the assessment of individual differences in deception production success. Additionally, the paradigm can easily be modified so that participants can choose to lie thus also allowing for the investigation of individual differences in tendency to lie.

Given these advantages the current study utilised DeceIT to investigate whether participants' tendency to lie and their success when doing so was associated with counterfactual thinking. Similar to our previous research we expect to find a positive association between these skills. Such results would imply that this relationship is robust and also relevant to situations which more closely resemble real-life events.

## Method

### Participants

A total of 98 of Plymouth University students took part in the study. Of these only 81 completed both the behavioural deception task and the counterfactual thinking measure, thus only these participants were included. Participants were 28 females and 53 males and ranged in age from 18 to 35 ($M$ = 20.62, $SD$ = 3.05). Participants were given course credit for their participation in the first part of the study and rewarded monetarily with 4£ for the second part of the study.

### Materials and Procedure

The experiment was completed in two separate sessions. In the first session participants completed the behavioural deception task, in the second session participants were administered the counterfactual thinking test.

**Behavioural deception task.** For the first session, participants were tested in either groups of five or four. Participants were told that they were to take part in a game designed to test their communication skills. As part of the game participants were required to make either truthful or false statements in front of the other members of the group. The aim was for the statements made to be as credible as possible. Credibility was decided by the other members of the group which had to judge whether the statements they heard were true or not. Participants took turns to either send messages (Sender role) or to judge other's messages (Receiver role). As our interest in this study is related to the ability to deceive and not the ability to detect deception, we were only interested in participants' performance in the Sender role.

Participants had to make statements in relation to their own past events. The experiment also included two additional topics (i.e. descriptions of emotionally-charged pictures and opinions about controversial subjects), which were not of interest to the current study as we were only concerned with deception in relation to episodic events.

The episodic memory phase consisted of 10 trials. On each trial, the participant was firstly presented with a cue which specified the topic, for example: 'Describe what you did last weekend'. All topics are presented in Appendix 3A. On three trials participants were instructed to lie, on three others they were asked to tell the truth and for the remaining four trials participants could choose whether to tell the truth or lie. Cues to lie, tell the truth or choose were presented to the Sender in a randomized order. These cues as well as the topics were presented on a computer monitor placed in front of the Sender but facing away from the Receivers, thus only the Sender had access to this information. Once the instructions were given, participants had 20 seconds to make their statement. The beginning and end of the time-frame was signalled by a loud beep. To make sure the Sender followed the cue correctly, and to ascertain their choice in the Choice trial, after each trial Senders were asked to confirm whether they told the truth or lied. This was done in hidden view from the Receiver by ticking one of two boxes on a spreadsheet that the Sender held in their hands. Simultaneously, Receivers also noted down whether they believed the statement given by the Sender was true or false. Response Sheets for both Sender and Receiver roles are presented in Appendix 3B.

In order to entice competitiveness, participants were told that they could win two prizes. One prize was awarded to the Sender who was most convincing as

measured by Receiver's scores and the other prize for the Receiver who was most accurate in their judgement as measured by how accurately they could detect lies and truths.

Performance in the Sender role was analysed using the sensitivity index d prime used in Signal Detection Theory (SDT) (Green & Swets, 1966). The index denotes the extent to which the Sender's lies and truths are believable in relation to one another has previously been used by Wright's et al. (2012) as an indicator of individual's ability to lie. Just as in Wright et al. (2012) study the following formula was used: z[P(Hit)]-z[P(False-Alarm)], where P(Hit) was the proportion of times a truthful message told by the Sender was correctly classified as a truth by Receivers and (P(False-Alarm) was the proportion of times a lie told by the Sender was incorrectly classified as a truth by the Receivers. The formula is in effect an index of the Sender's success when lying relative to success when telling the truth. A positive score indicated that a Sender was more successful when being truthful than when lying (i.e. truths were believed more than lies), a score of 0 indicated that the Senders' performance was equivalent both when telling the truth and when lying (both lies, and truths were equally believed). Lastly, a negative score was indicative of the fact a Sender was more successful when telling a lie than when telling the truth (their lies were more plausible than their truths). Wright et al. (2012) interpreted that individuals with a negative score can produce lies that are less detectable and thus understood this to be an indicator of lie production success, in this thesis we will adopt the same interpretation.

**Counterfactual thinking measure.** To measure spontaneous and cued counterfactual thoughts we used the 'moving town' scenario used in our previous

studies (McEleney & Byrne, 2006). Participants were given 5 minutes to write a free narrative of their thoughts and feelings in the style of a personal diary entry and we counted the number of counterfactuals generated in each narrative. We used the same coding scheme as in our previous studies.

## Results

Seven participants did not correctly follow the cue given in one trial. Four of the participants told the truth when instructed to lie, the other three told a lie when cued to tell the truth. For these participants performance in the Sender role was calculated based on the remainder nine trials in which the cue was followed correctly.

### Association between counterfactuals and the number of lies in the Choice trails

We did not find a statistically significant association between spontaneous counterfactual thinking and the number of deceptive statements participants made when given the choice to do so ($r$ =-.01; $p$ = .984). The number of cued counterfactuals also did not correlated with the choice to lie ($r$ = -.13; $p$ = .256).

### Counterfactual thinking and lie production success

As predicted, participants who came up with more spontaneous counterfactuals on the diary tasks were more likely to be successful when lying than when telling the truth ($r$ = -.26, $p$ = .019). The number of cued counterfactuals did not correlate with lie production success ($r$ = -.14, $p$ = .221).

## Discussion Study 6

Deception production skill was positively associated with spontaneous counterfactual thinking. Our previous studies only investigated the tendency to lie rather than the successful ability to do so, therefore these results present new insights into the relationship between counterfactual thinking and deception. These results are also consistent with some of the findings in Study 5 which suggest that lying is easier and more credible when told following a counterfactually heightened situation.

The fact that counterfactual thinkers were more successful when producing lies could be explained by several factors. Firstly, when attempting to detect lies receivers have been found to rely partly on non-verbal cues (Vrij, Granhag, & Mann, 2010). It could be hypothesised that Senders which are experienced in generating alternatives to reality can more easily come up with a plausible statement. This might in turn imply that less cognitive resources need to be allocated to this task, permitting the allocation of cognitive resources to the monitoring and inhibition of non-verbal cues.

Secondly, in the DeceIT paradigm, Receivers have also been found to rely on response latency in order to judge whether a statement is truthful or deceptive (Wright et al., 2012). That is, Senders who are quicker at making deceptive statements tend to be judged as more successful. It could be that people who can easily engage in counterfactual thinking are quicker when deceiving because alternatives to reality are more readily available to them.

Besides nonverbal cues, another potential explanation is that the lies generated by counterfactual thinkers are less detectable because of their content. Lies

that differ only slightly from the truth are more difficult to detect (Vrij et al., 2010). Thus, it could be that people who find it easier to generate alternatives also find it easier to come up with a more plausible lie. All of these hypotheses are given tentatively and should be tested directly in future research.

In this study we found that participants who spontaneously came up with more counterfactuals did not choose to lie more when given the option. This is in contrast with our findings in Study 1 and our interpretation that counterfactuals play a role in deception through a counterfactual-neutral pathway. Two explanations could clarify this discrepancy. Firstly, the DeceIT paradigm introduces an element of competitiveness in order to motivate participants to succeed in their task. The motivation of a prize should have enabled Senders to strive to be as successful as possible, namely to appear as truthful when telling the truth and when lying. However, when presented with a choice, Sender might have reasoned that telling the truth comes more naturally and as such the probability of being successful whilst being truthful is higher than the probability of being successful whilst lying. In order words, Senders might have concluded that the benefits of choosing to tell the truth were higher. Thus, the competitive nature of this game might have motivated the construction of successful lies but not the decision to deceive. Indeed, a higher percentage of participants chose to be consistently truthful (i.e. never lied when given the choice) in the current study (33.3%) as opposed to consistently truthful in Study 1 (i.e. never lied on either scenario in response to spontaneous deception questions) (10.3%)[5].

[5] The percentage of truthful responses for this study was calculated across all four Choice trials. For Study 1 this percentage was calculated by taking into account responses to both scenarios.

Secondly, the current experiment also differed from our previous studies in regards to the relationship between the deception task and counterfactual thinking task. In the current study, the counterfactual task was liked to a hypothetical scenario whereas the deception was linked to a past episodic event. It is possible that in order to notice association participants should also experience the counterfactual nature of the event which they then have to falsify or represent truthfully. This is also important given the evidence we noted earlier about how different perspectives allow for the differential construction of counterfactual thoughts (Girotto et al., 2007). It is therefore necessary to assess the link between deception and counterfactuals with a behavioural task that can concurrently assess both counterfactual thinking and deception. Study 7 aims to utilize such a paradigm in order to understand whether individuals use the information available through counterfactual thinking in order to lie.

## Study 7

The current study was developed in order to further ascertain whether people's tendency to lie is motivated by the counterfactual nature of a situation. We aimed to design a novel behavioural paradigm in which we could directly manipulate counterfactual alternatives and where we could directly assess decisions to lie at an individual level. Based on the deception decision task developed by Gneezy, Rockenbach, and Serra-Garcia (2013), we designed a game in which participants could lie in order to increase their own benefits at no costs to themselves but potentially at a cost for another. The fact that participants could gain concrete rewards by lying, meant that the motivation to choose to lie was increased compared to our previous studies. Furthermore, the fact that the decision to lie could affect

someone else meant that this decision had consequences. To manipulate counterfactual availability, participants played a game with multiple trials in which they lost by a narrow or wide margin. If deception is influenced by thoughts of how things could have happened differently, people should lie more often in instances where the desired outcome was narrowly missed than in situations where the desired outcome was completely missed.

## Method

### Participants

Sixty one participants took part in this study (53 females and 8 males). Participants ranged in age from 18 to 50 years ($M$ = 21.31, $SD$ = 6.62). All participants were undergraduate students at Plymouth University who received course credit for their participation.

### Materials and Procedure

Participants had 24 trials on a slot machine game and their aim was to accumulate as many points as possible. The slot machine had two columns of spinning numbers. Participants pressed a button to stop the numbers spinning and they won if the sum of the numbers matched a target number (ranging from 1-8) which was presented at the top of the screen (see figure 3.1).

There were three types of trials: win, near-miss and loss and each participant received an equal number of trials per type, (i.e. 8 goes in each of the win, near-miss and loss conditions). In the *win* condition, the slot machine always stopped on two numbers which correctly added up to the target number. In the *near-miss* condition, one column stopped between two numbers. Participants were told if this happened,

then the lower number would count. In all cases, if the higher number was added to the second number, then the total would have matched the target (see figure 3.1 for an example). In the *loss* condition, one of the numbers on the slot machine was higher than the target number by one point and the sum of the two numbers was greater than the target by 2 or 3. For similarity with the near-miss condition, on half of the trials one of the columns on the slot machine stopped in between two numbers.

At the end of each trial, the sum of the two numbers was displayed on the screen (see figure 3.1). In addition, the screen turned green if they won and red if they lost. Participants were then presented with a screen asking them to report whether they had won. They clicked ‘Y’ for yes and ‘N’ for no and they received one point each time that they clicked ‘Y’, regardless of whether they actually won on that trial.

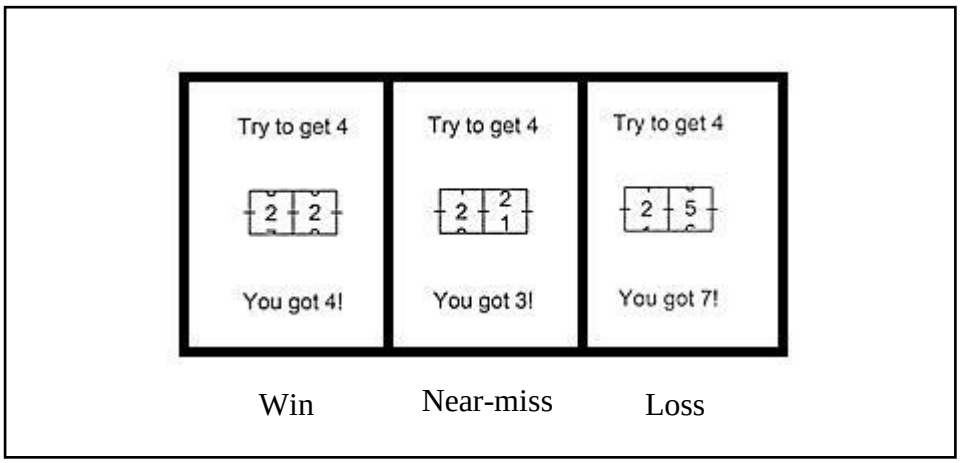


*Figure 3.1*. Pre-determined numbers in each condition for target number 4.

The study was carried out in small groups and participants were told it involved a multi-player game. They were competing against a fictitious player B and the one who accumulated the most points in the slot machine game would win a prize of 15 minutes of participation credit. To encourage participants to tell the truth, they were also told that the game included a third player, C, who could win a prize by judging whether the other players were truthful about their reported wins. They were told that to help participant C, they should tell the truth because previous studies show that this is what they are likely to assume. The full instructions are presented in Appendix 3C. To enhance belief in the two fictitious players, participants had a chance to send a short message to them before the game began. In return they also received two messages. The message from participant B read: 'Good luck, but I will try to win this', and the message from participant C read: 'Please let me win, be truthful'. The game commenced after the two messages were displayed.

## Results

Overall, 19 participants (31.1%), were truthful on all trials, and only 8 (13.1%) lied on all trials in which they would have gained from lying. The remaining 34 participants (55.7%) lied on some trials but not others. Most importantly, as predicted, participants lied significantly more in the near-miss condition ($M$ = 3.21, $SD$ = 3.10) than in the loss condition ($M$ = 2.77, $SD$ = 2.83), $t(60) = -2.30$, $p = .020$, $d = .31$. The result is consistent with our earlier findings which show that counterfactual availability increases the accessibility of lies and in addition shows that individuals will use these lies when given the opportunity.

## Discussion Study 7

Similar to past research only a limited number of participants engaged in maximal possible dishonesty, i.e. lied on all trials (Mazar, Amir, & Ariely, 2008). When participants chose to lie they were more likely to do so in situations where counterfactual alternatives were highly available. These findings highlight that when the motivation to lie is high and the counterfactual task is linked to the situation, people make use of the counterfactual nature of a situation in order to deceive. These findings support our findings from Chapter 2 as well as previous research concerned with the malleable nature of our ethical principles (Moore & Gino, 2013).

As mentioned previously we believe that a range of mechanisms could underlie this relationship. This is a novel paradigm and future studies could seek to utilise this behavioural research paradigm in order to elucidate the mechanisms underlying this relationship.

## General Discussion

Overall the studies in this chapter support our previous results and suggest that the positive relationship between counterfactual thinking and deception is maintained even when tested with measures which more closely resemble the demands of real-life situations.

The results in this chapter also add new insight by revealing that people who are more likely to engage in counterfactual thinking tell lies that are more difficult to detect. This could have potential implication for the development of lie detection techniques which should be calibrated in order to account for such individual

differences. Nevertheless, future work should be done in order to replicate these findings and further investigate the factors which underscore this link.

The paradigms of deception we have used in this chapter improve on the methods previously used, however they do not address a very much debated limitation of deception research, namely the fact that lies were sanctioned. In Study 7 for example it was made clear to participants that they could lie and that this would be permitted by the experimenter. Permitting participants to deceive differs from real-life situations where such communication is usually frowned upon. Some have argued that sanctioned lies are less cognitively demanding and less affectively charged because they make liars feel less guilty and less motivated than what would normally be the case (Ekman, 1988; Vrij, 2000). However, others have found little or no differences in the experimental investigation of sanctioned versus unsanctioned lies (Feeley, 1996; Sporer & Schwandt, 2007). Additionally, the lies in this study were not directed towards the experimenter but rather towards a third party who would not have sanctioned such behaviour. Withstanding this limitation the current chapter provides us further evidence that the imagination of counterfactual alternatives can influence deceptive behaviour.

# CHAPTER 4

## Does counterfactually derived affect impact deceptive communication?

Findings from previous chapters established a positive relationship between counterfactual thinking and deceptive communication both when using hypothetical scenarios and when using behavioural measures. In the studies to follow we wish to investigate this relationship further by examining its underpinnings. In particular, this chapter aims to assess whether the affective consequences of counterfactual thinking play a role in its relationship with deception.

In previous chapters we have suggested that counterfactuals influence deception though a content neutral pathway. We reasoned that because the content of counterfactual thoughts did not directly transfer to the content of lies, counterfactuals might facilitate deception by providing general information that things could have been different. However, another possible interpretation which also aligns with the content neutral pathway is that counterfactuals might influence dishonesty by changing the way people feel in relation to a past event. Some results from Study 5 suggest that when counterfactual alternatives are readily available an individual might feel more negative affect which in turn motivates them to deceive. For example, some participants reasoned that missing a postal delivery by only one minute as opposed to 15, Ed would be more likely to lie because he would ‘feel more guilty as he could have gotten there on time if he went faster’. In order to further investigate this potential link, the current chapter aims to explore whether the way that counterfactuals make us feel in relation to an outcome can impact subsequent deceptive communication. An examination of whether emotional factors underpin

our previous results will allow us to understand whether the mechanisms involved are partly motivational.

Affect has been extensively investigated in connection to counterfactual thinking. Contrasting ‘what could have been’ with ‘what is’ can elicit diverse emotional reactions. Affect can be amplified or attenuated depending on the degree of similarity between the actual outcome and the imagined counterfactual outcome (Gleicher et al., 1990). If the imagined counterfactual possibility leads to an outcome conflicting with the actual event, affect will be amplified. However, if the counterfactually imagined alternative leads to the same outcome as the actual event, affect will be minimised.

The direction of the counterfactual thought is important in determining the type of affective consequences. Most research has shown that if after experiencing a negative event we engage in upward counterfactual thoughts, (i.e. thoughts about how things could have been better), this can lead to negative feelings such as feelings of dissatisfaction. Conversely, downward counterfactual thoughts which involve imagining how things could have been worse can lead to the opposite affective response, i.e. feelings of relief (Markman, Gavanski, Sherman, & Mcmullen, 1993; Medvec, Gilovich, & Madey, 1995; Roese, 1994).

Counterfactuals which are constructed to evaluate the outcome of an event are also known to amplify and shape emotions relevant to morality such as feelings of guilt and shame. Niedenthal et al. (1994) showed that constructing counterfactual alternatives which involve changing details related to one’s own traits or to one’s behaviour can differentially either lead to feelings of shame or guilt respectively. For example, undoing a negative situation by thinking ‘If only I was more thoughtful’

could lead to feelings of shame, whereas constructing a counterfactual thought such as ‘If only I had not eaten the last piece of cake’ can lead to guilt. In related findings, Zeelenberg et al. (1998) found that behaviour focused counterfactual thoughts can lead to regret, whereas counterfactual thoughts which focus on how aspects of the situation could have been different lead to feelings of disappointment. Taken together this suggests that counterfactual thoughts are closely linked to affect and that different emotional states can follow counterfactual thoughts depending on factors such as direction, control and degree of comparison.

Affective responses are also an important catalyst for ethical judgements and behaviours. General negative affect can impact deceptive behaviour, for example, anxiety, characterised by unpleasant negative emotional states, has been shown to impact ethical judgements and behaviours. Kouchaki and Desai (2015) found that anxious individuals are more likely to be perceived as being dishonest and more willing to participate in unethical acts such as cheating in order to gain monetary rewards. We were not able to replicate these findings, as in Study 1 we did not find any association between anxiety and the number of lies participants generated. However this could potentially be because we measured trait anxiety whereas Kouchaki and Desai (2015) assessed state anxiety.

The emotional factors associated with rejection can also entice people to deceive more. Van Der Zee et al. (2016) showed that participants who had an imaginary insurance claim rejected were more likely to falsify information in a following claim. Results also revealed that a reduction in happiness following rejection could accurately predict the relationship between rejection and the increase in deceptive behaviour. Moreover, other studies have shown that guilt following a

misdeed (i.e. post-transgressional guilt) can be associated with an increase in deception in a subsequent unrelated task (Depalma et al., 1995). This is in line with our own findings from study 5 with suggesting that guilt could motivate lying.

However, it also needs to be mentioned that some studies have found a negative relationship between negative emotions such as guilt and unethical behaviour. For example, an aversion towards guilt or shame can lead to a preference for truth telling (Greenberg, Smeets, & Zhurakhovska, 2014). Similarly, increased feelings of guilt can also lead to confessions of marital infidelity (Baumeister, Stillwell, & Heatherton, 1995). Others also suggest that the more intense the negative feelings following a transgression, the more one might be motivated to fully confess in order to minimize these feelings (Peer, Acquisti, & Shalvi, 2014).

Overall this evidence suggests that counterfactuals can amplify negative affect and this negative affect can at times motivate deception, but this might depend on the characteristics of a situation, and the intensity of the negative feelings. Despite this potential link, no studies have so far directly investigated affect as a potential mechanism underlying the relationship between counterfactuals and deception. Shalvi et al. (2011) have mentioned that the observation of a desired counterfactual might decrease the negative feelings usually associated with lying by emphasizing that the lie to be told in such situations is minor. However, this interpretation relates to the way in which counterfactuals make us feel about the lie rather than the outcome of the event itself. Gaspar et al. (2015) have also shown that engaging in counterfactual considerations following a misdeed can lead to feelings of guilt which in turn reduce the tendency to be deceptive on a subsequent unrelated task. However, in this study guilt was assessed independent of the decision to engage

in deceptive behaviour and therefore this link is not a direct one. Furthermore, it is still unclear how guilt following counterfactual thinking might impact lying about the same event. Therefore, the question concerning the impact of counterfactually derived affect on subsequent deception remains to be answered.

In order to address this gap in the literature, the studies in this chapter will aim to test the way in which general affect mediated by downward and upward counterfactual thoughts influence subsequent deception (Study 8 and Study 9). In addition, we will also be looking at whether moral affect derived from counterfactual thoughts about perceived feelings of control can impact subsequent tendency to lie (Study 10). Given that in our previous studies we did not find the experimental manipulation of counterfactual alternatives to influence cued generations, the current chapter focuses on spontaneous generations alone.

## Study 8

The current study aims to investigate whether the affect derived from counterfactuals thoughts meant to evaluate the outcome of a negative event can impact subsequent deceptive communication. As previously mentioned, the direction of evaluative counterfactual thoughts has been shown to influence subsequent affective reactions. Compared to a no-counterfactual condition, upward counterfactuals (thinking about how things could have been better) are more powerful at promoting negative affect whereas downward counterfactuals (thinking about how things could have been worse) tend to promote positive affect (Markman et al., 1993; Roese, 1994). We aimed to use the direction of counterfactual thoughts in order to manipulate affective reactions and examine their influence on subsequent deception. We manipulated counterfactual direction by asking participants to engage

in either factual, upward or downward thoughts by completing sentence stems in line with a requested direction. We expect people in the downward condition to report feeling the best and people in the upward condition to report feeling the worst.

Consequently, if the relationship between counterfactuals and deceptive communication is influenced by negative emotion, we hypothesise that the quantity of deceptive responses will differ between experimental conditions in a way which follows the pattern of affective responses. Namely we expect that people in the upward condition will be motivated to lie the most whereas people in the downward condition will lie less. The way in which affect impacts deception might be dependent on the characteristics of a situation, as such we also included scenario type as a separate variable in the analysis.

The difference between scenarios has not previously been investigated in this thesis because the focus in previous chapters has been different. In Chapter 2 and 3 we collapsed responses across scenarios because we were interested in the relationship between counterfactuals and deception in general (i.e. notwithstanding the characteristics of a situation). The focus on the current chapter calls for a more thorough investigation of the effect of individual scenarios because each scenario refers to a different type of rule violation which is known to influence affective reactions differently (Wilson & O'Gorman, 2003).

## Method

### Participants

In total, 151 participants participated in the online study and completed all questions. Participants who did not answer all questions were excluded. Six participants allocated to the factual condition were excluded because they provided

counterfactual thoughts in their diary entry (two participants for each of the three scenarios). Six other participants were excluded because they shared IP-addresses. The remaining 139 (82 females and 57 males) which were included in the study ranged in age from 18 to 61 (*M* = 25.73, *SD* = 7.78). Participants were recruited and performed the study online and were not rewarded for their participation..

## Materials and Procedure

**Counterfactual manipulation.** We used a shorter version of the 'moving town' and 'car incident' scenarios used in Study 1. We also used a third scenario which was adapted from Scenario 2 in Study 4 and described a protagonist missing a housemate's postal delivery due to unexpectedly stopping for a coffee on the way home. This scenario was chosen because it had the most affect related responses in Study 5. All three scenarios are presented in Appendix 4A. Due to the length of the test battery, and thus in order to avoid high attrition rates, participants were randomly assigned to only one of three scenarios ( i.e. moving town, car incident or delivery) and only one of three thought conditions (i.e. factual, downward or upward).

In each condition participants were first instructed to read the scenario they had been assigned to. After reading the scenario, depending on the condition they were allocated to, participants were asked to vividly imagine the events as if they had happened to them and to imagine writing a diary page about them. Participants were then given sentence stems to complete. For the factual condition participants were asked to reflect on how things ultimately turned out and they were then provided with sentence stems which emphasised the outcome of the scenario, i.e. 'Today I missed my neighbour's party…' (moving town scenario) and 'Today I hit

another car…' (car incident scenario) and 'Today I missed my housemate's delivery…' (delivery scenario). Participants were asked provide two possible endings.

In the upward condition at the end of the scenarios participants were asked to reflect on how things could have turned out better. They were provided with the same sentence stems as in the factual condition, but this time each sentence stem also read: 'Things might have turned out better if…..'. Participants had to complete this sentence twice.

Finally, in the downward condition participants were asked to reflect back on the events described in the scenario and imagine how things could have turned out worse. The factual sentence stems were also accompanied by the following sentence: 'Things might have turned out worse if…..'. Participants had to complete this sentence stem twice.

**Measure of affect.** After completing the sentence stems, to ensure that counterfactual thoughts were constructed to evaluate the information in the scenario we asked participants to consider how these thoughts would make them feel about the events described in the scenario. Specifically they were asked: '*How would the thoughts you generated above make you feel in relation to the events described above?*' They had to do this using a using a 7 point scale on which had 1 indicated 'very good' and 7 indicated 'very bad'. Higher scores indicated increased negative affect.

**Ease of generation measure.** Because upward counterfactuals are easier to generate than downward counterfactuals (Markman et al., 1993) we also wanted to record the ease with which participants complete the sentence stems. Participants

were also asked to rate the ease of generation on a scale from 1 to 7 on which 1 was labelled as very easy and 7 as very difficult.

**Measure of deception.** In order to assess spontaneous deception and in line with our previous studies, participants were instructed to vividly imagine either bumping into their neighbour (moving town scenario), being asked questions by the police (car incident scenario) or talking to their housemate (delivery scenario). They read: *'Several possible answers come to your mind immediately. Please write down some of the different things you would say to your neighbour/ police/ housemate in reply.'* They were provided with 5 spaces in which they could enter free text. This differed from previous studies where participants had been provided with a single blank space. The aim was to avoid potential interpretation bias. In previous studies not all participants might have been aware that they could provide more than one potential answer, thus we wanted to control for this by making sure participants could interpret that more than one answer could be entered.

Similar to previous studies in this thesis, lies were coded according to the definition provided by Serota and colleagues (2010) and the coding scheme attached in Appendix 2D. Each individual lie was given a score of 1.

## Results

The assumption of normality for the outcome variables was violated. Results were analysed using both parametric and non-parametric method. For the non-parametric analysis we use the Scheirer-Ray-Hare extension of the Kruskal-Wallis test, for the parametric analysis we used a two-way ANOVA analysis. The patterns

of results were similar for both analyses. For ease of interpretation however we chose to report the parametric analysis here.

**Affect results**

Descriptive results for rating of affect for each scenario and each condition are provided in Table 4.1. A two-way ANOVA was conducted in order to examine the effect of condition (factual vs upward vs downward) and scenario ('moving town' vs 'car incident' vs 'delivery') on ratings of affect. The results revealed no significant main effect of Condition, $F$ (2, 139) = 1.40, $p$ = .250, $\eta_p^2$ = .02, a significant main effect for Scenario, $F$ (2, 139) = 5.17, $p$ = .007, $\eta p^2$ = .07 and no interaction between scenario and condition, $F$ (4, 139) = 1.91, $p$ = .113, $\eta p^2$ = .06. A Tukey post hoc tests revealed that participants allocated to the car crash scenario felt particularly worse in comparison to those allocated to the moving town scenario ($p$ = .006), however, there was no difference between the car crash and delivery scenario ($p$ = .568), or the delivery scenario and the moving town scenario ($p$ = .124).

Table 4.1.

*Means and standard deviations for ratings of affect for each scenario across experimental conditions*

| Scenario | Condition | | |
|---|---|---|---|
| | Factual | Upward | Downward |
| | *M* (*SD*) | *M* (*SD*) | *M* (*SD*) |
| Moving town | 4.88 (1.36) | 4.95 (1.40) | 5.31 (0.95) |
| Car crash | 6.36 (1.08) | 5.11 (1.81) | 6.13 (1.19) |
| Delivery | 5.36 (1.03) | 5.67 (0.82) | 5.53 (0.99) |
| Totals | 5.50 (1.33) | 5.20 (1.43) | 5.67 (1.09) |

*Note:* Higher ratings indicate greater negative affect.

**Ease of generating counterfactual statements**

To assess whether participants found it differentially difficult to generate the counterfactual statements, we performed a two way ANOVA with scenario ('moving town' vs 'car incident' vs 'delivery') and counterfactual condition (factual vs upward vs downward) as independent variables and ease of generation as the dependant variable. The results revealed a significant main effect of counterfactual condition, $F$ (2, 139) = 14.30, $p<.001$, $\eta_p^2$ = .18, no significant main effect for scenario type $F$ (2, 139) = .83, $p$ = .439, $\eta_p^2$ = .01 and a significant interaction between scenario type and counterfactual condition $F$ (4, 139) = 3.59, $p$ = .008, $\eta_p^2$ = .10. Descriptive statistics are presented in Table 4.2.

Table 4.2.

*Means and standard deviations for ease of generating counterfactuals for each scenario across experimental conditions*

| Scenario | Condition | | |
|---|---|---|---|
| | Factual | Upward | Downward |
| | *M* (*SD*) | *M* (*SD*) | *M* (*SD*) |
| Moving town | 2.64 (1.62) | 1.86 (0.83) | 3.46 (1.80) |
| Car crash | 2.86 (1.79) | 2.50 (1.58) | 2.88 (1.45) |
| Delivery | 2.21 (1.19) | 2.13 (1.31) | 4.93 (1.79) |
| Totals | 2.58 (1.54) | 2.14 (1.26) | 3.75 (1.86) |

*Note:* Higher ratings indicate greater difficulty in generating a counterfactual statement.

Simple main effects analyses were used in order to follow up the interaction. The mean comparison as a function of scenario type revealed that there were no

significant differences between the factual and upward condition for either the moving town or car incident scenarios (all *p's* >.132). For the moving town scenario coming up with a counterfactual thought was more difficult in the downward condition ($p$ = .004) as compared to the upward condition. This difference was not significant for the car incident scenario ($p$ = .614). For the delivery scenario, those in the downward condition found it significantly more difficult to come up with a counterfactual thought in comparison to both other conditions (both *p's* <.001), but there were no differences between participants in the factual condition and the upward condition ($p$ = .855).

**Spontaneous deception results**

Descriptive statistics for spontaneous deception are provided in Table 4.3. We performed a two way ANOVA in order to assess the influence of scenario type ('moving town' vs 'car incident' vs 'delivery') and counterfactual condition (factual vs upward vs downward) on the number of spontaneous lies participants generated.

The results revealed a significant main effect for counterfactual condition $F$ (2, 139) = 3.12, $p$ = .048, $\eta_p^2$ = .06, no significant main effect of scenario, $F$ (2, 139) = 1.95, $p$ = .146, $\eta_p^2$ = .03, and no interaction between scenario and counterfactual condition $F$ (4, 139) = 0.26, $p$ = .902, $\eta_p^2$ = .01. A Tukey post hoc tests revealed that participants allocated to the downward condition told significantly more lies comparison to those allocated to the factual condition ($p$ = .050), however, there was no difference between the upward condition and the downward condition ($p$ = .397), or between the factual and upward condition ($p$ = .185).

Table 4.3.

*Means and standard deviations for spontaneous lies for each scenario across experimental conditions*

| Scenario | Condition | | |
|---|---|---|---|
| | Factual<br>*M* (*SD*) | Upward<br>*M* (*SD*) | Downward<br>*M* (*SD*) |
| Moving town | 2.24 (1.35) | 2.62 (1.88) | 3.00 (1.96) |
| Car crash | 1.50 (1.29) | 2.17 (1.30) | 2.80 (1.42) |
| Delivery | 1.82 (1.78) | 1.80 (1.42) | 2.33 (1.76) |
| Totals | 1.88 (1.45) | 2.24 (1.59) | 2.70 (1.61) |

We also ran the same analysis with affect as a covariate. This did not change the pattern of results as described above. The main effect of scenario was again not significant ($p$ = .179), the main effect of condition remained significant ($p$ = .046), and there was no significant interaction between condition and scenario ($p$ = .783)

Given that participants distributed to the upward condition differed in terms of how easy they found it to complete the sentence stems depending on the scenario they allocated to, we also wanted to perform the analysis with ease of generating counterfactuals as a covariate. The main effect of scenario remained not significant ($p$ = .126), additionally, the main effect of counterfactual condition also disappeared ($p$ = .124). Similarly, there was no significant interaction between scenario type and counterfactual condition ($p$ = .841). However, in the downward condition, the ease of generation score did not correlate with the number of lies that participants generated ($r$ = .03, $p$ = .830).

## Discussion Study 8

We did not find any differences in terms of affect between any of the three counterfactual conditions and thus the experimental manipulation was not successful. These results are in contrast with research which suggests that the direction of counterfactual thoughts can promote differential affect (Markman et al., 1993; McMullen, 1997; McMullen, Markman, & Gavanski, 1995). A potential reason for why we could not replicate such previous findings could be related to the fact that affect was only measured using one item. Single-item scales can be more vulnerable to random measurement error such as unknown biases in interpretation (Credé, Harms, Niehorster, & Gaye-Valentine, 2012). As such it could be that the outcome measure for affect used in this study did not allow participants to express exactly how the counterfactual thoughts they generated would make them feel. This limitation could be addressed by using multiple-item scales which can better account for such limitations whilst also covering a wider range of the affect construct.

The only difference in affect identified was the one between the two scenarios used in our previous studies. Following the creation of counterfactual thoughts participants allocated to the 'car incident' scenario reported feeling significantly worse than those allocated to the 'moving town' scenario. This result is congruent with findings which suggest that the transgression of 'weak norms' (such as breaking the promise to a neighbour as in the 'moving town scenario') is associated with a reduced emotional reaction in comparison to the transgression of 'strong norms' (such as hitting a car in the 'car crash' scenario) (Wilson & O'Gorman, 2003). This was to be expected given that the reason we included these two scenarios was because they involve different levels of risk for the liar.

In terms of deception, participants who generated thoughts about how things could have been worse (i.e. downward condition) generated more lies than those who had to think about the characteristics of the event (factual condition). However, these results did not appear to be based on counterfactually derived affect. It could be that this difference is due to a potential variance in the availability of information across conditions. Participants in the downward condition found it more difficult to come up with counterfactual thoughts in comparison to participants in the other two conditions. This aligns with previous findings which have shown that we are more likely to come up with upward counterfactuals rather than downward ones (Roese and Olsen, 1997). This potentially suggests that participants in the downward condition lied in comparison to the factual condition more because they had to mentally generate a wider range of alternatives before reaching an appropriate response. It is also possible that whilst searching for ways in which things could have been worse participants also thought about several ways in which the event could have been better. These participants could have therefore been more aware about how things could have been different. The difference in deception between the downward and factual condition was not significant once controlling for ease of generation ratings. This interpretation might however be contentious given that ease of generating counterfactuals in the downward condition was not related to the number of deceptive responses. The availability of information interpretation might also be questionable given that the number of lies was not different between the factual and upward condition. However this might be explained as our experimental manipulation might not have stopped people in the factual condition from generating counterfactual thoughts spontaneously.

Overall then, differences in the generation of deceptive statements could not be accounted but the way in which the counterfactuals made participants feel about the outcome of the event.

However, given that the experimental manipulation did not work, several limitations of the study would need to be addressed before we can draw a conclusion about counterfactually derived affect and deception. Firstly, as already suggested above, affect should be measured using a multiple item scale. Secondly, differences between scenarios might have been diminished as the scenarios were administered between participants rather than within. Interpretation could be enhanced if participants were assigned to one experimental condition but provide them with a chance to respond to all scenarios.

Additionally, if we consider that the same type of counterfactual thought can lead to different affect depending on which factors one concentrates on (McMullen, 1997; Niedenthal et al., 1994) it could also be that there were differences between participants within each condition. Because participants created their own counterfactual thoughts they could have concentrated on entirely different things. Indeed, we find such differences when looking at the completed sentence stems. For example, when providing downward counterfactual thoughts for the car crash scenario some participant's thoughts were self-directed, such as 'It would have been worse for me if someone had seen me' but others concentrated on other-directed elements, such as 'It would have been worse if I had hit somebody'. One way to address this would be to control for the specific thoughts that participant are asked to imagine by providing participants with the thought rather than allowing them to

created it themselves. We aimed to address these limitations in the study which follows.

## Study 9

The manipulation in our previous study was not successful. The current study was designed in order to address the limitation of study 8. First we aimed to directly provide participants with factual, upward or downward thoughts in order to avoid differences within each condition and to eliminate differential generation ease between conditions. Secondly, we also aimed to assess affect using a multi-item scale, which we hoped would provide with a wider measurement for the construct of negative affect. If negative affect resulting from counterfactual thoughts influences the likelihood that people lie, then we expect that deception will be most frequent in the downward condition.

## Method

### Participants

A total of 149 (127 female and 22 male) participants took part in the study. Participants were Plymouth University undergraduate students who received course credit for their participation. They ranged in age from 18 to 43 years ($M$ = 21.41, $SD$ = 5.17). Participants took part in the study outside of the lab by accessing a web link.

### Materials and procedure

Given that in Study 8, participants found it considerably more difficult to imagine counterfactuals for the delivery scenario, the current study used only the ‘moving town’ scenario and the ‘car incident’ scenario. Consequently, the battery

was more suitable for online testing and allowed participants to respond to both scenarios.

**Counterfactual manipulation.** Participants were randomly assigned to one of three conditions factual, downward or upward. In each condition participants were first instructed to read the two scenarios, i.e. the ‘moving town’ scenario and the ‘car incident’ scenario. Participants read each scenario at a time and answered questions relating to that scenario before moving on to the second scenario. Scenarios were presented in a counterbalanced order. After reading the scenario depending on the experimental condition they were assigned to, participants were asked to read a sentence. For the factual condition participants were simply asked to reflect on how the events in the scenario ultimately turned out and they were instructed to imagine thinking about the outcome of the events. For the ‘moving town’ scenario they were told to think *‘Today I missed my neighbour's party in order to attend a movie on my own.’* For the ‘car incident’ scenario they were asked to think *‘Today I hit a car’*.

In the upward condition at the end of the scenarios participants were asked to reflect on how things could have turned out better and were provided with an example of such a thought. For the ‘moving town’ scenario the sentence read *‘Things might have turned out better if I had not gone to the movie on my own'* whereas for the ‘car incident’ scenario they read *‘Things might have been better if I had not hit the car.’*

Finally in the downward condition participants were asked to reflect back on the events described in the scenario and imagine how things could have turned out worse. For the ‘moving town’ scenario the sentence read *‘Things might have turned*

*out worse, at least I didn't stay at home by myself*' whereas for the 'car incident' scenario they read *'Things might have been worse, at least I didn't hit an animal'*.

**Measure of affect.** After reading the scenarios and the sentences participants were asked to vividly imagine how thinking the counterfactual thoughts made them feel in relation to the events described in the scenario. For example in the factual condition participants were asked: *'How would this thought about how things ultimately turned out make you feel in relation to the events of the da*y? They were asked to report their affect on a 4 item Likert type scale on which options for each item ranged from 1 to 7. The ends of the continuum for each item were: good/bad, positive/ negative, happy/sad, relieved/ worried, respectively. Higher scores indicated more negative feelings. The order of the items was randomised.

**Deception measures.** In order to assess spontaneous deception, participants were instructed to vividly imagine either bumping into their neighbour (moving town scenario) or being questioned by the police (car incident scenario). They read: *'Several possible answers come to your mind immediately. Please write down some of the different things you can imagine saying to your neighbour/ police in reply'.* As before, they were provided with 5 spaces in which they could enter free text. Lies were coded based on the same definition of deception used in previous studies (Serota et al., 2010) and the coding scheme attached in Appendix 2D. Each individual lie was given a score of 1.

## Results

The assumption of normality was violated for the outcome variables. Results were analysed using both parametric and non-parametric methods. The non-

parametric analysis was performed using Generalized Estimating Equations (GEE) whereas the parametric analysis as performed using a mixed ANOVA. Pattern of results were similar for both analyses. For ease of interpretation however we chose to only report the parametric analysis here.

**Affect results**

A total affect score for each scenario was computed by adding together the results from each of the four affect item. The resulting total affect score could range from 4 to 28, with higher scores indicating more negative affect. We then performed a mixed ANOVA with the affect score for each scenario (moving town, car incident) as the within-subjects factor and condition (factual, upward, and downward) as the between-subjects factor. Descriptive statistics for affective responses to each scenario for each of the three conditions are presented in table 4.4.

Table 4.4.

*Means and standard deviations for ratings of affect in response to each scenario in each of the three experimental conditions*

| Scenario | Condition | | | Total |
|---|---|---|---|---|
| | Factual *M* (*SD*) | Upward *M* (*SD*) | Downward *M* (*SD*) | |
| Moving town | 20.26 (3.76) | 20.76 (3.39) | 16.98 (3.49) | 19.35 (3.87) |
| Car crash | 21.56 (4.35) | 22.06 (4.31) | 17.41 (4.48) | 20.36 (4.82) |
| Totals | 20.91 (2.81) | 21.41 (3.19) | 17.19 (2.59) | - |

Analysis revealed that there was a significant main effect of scenario $F$ (1, 147) = 5.08, $p$ =.026, $\eta_p^2$ =.03 with participants responding more negatively to the

‘car incident’ scenario than the ‘moving town’ scenario. There was also a significant main effect for condition $F$ (2, 147) = 31.62, $p < .001$, $\eta_p^2 = .30$ but no significant interaction between scenario and condition $F$ (2, 147) = 0.42, $p$ =.660, $\eta_p^2 = .01$. A Tukey post hoc tests revealed that participants in the downward condition felt significantly more positive when compared to those in the upward condition ($p < .001$) and factual condition ($p < .001$). However there were no significant differences in terms of affect between participants in the factual condition and those in the upward condition ($p = .661$).

**Spontaneous deception results**

In terms of spontaneous deception, a mixed ANOVA with the number of lies generated in response to each scenario (moving town, car incident) as the within-subjects factor and condition (factual, upward, downward) as the between-subjects factor revealed a significant main effect for scenario $F$ (1, 147) = 14.39, $p < .001$, $\eta_p^2 =.09$, with participants coming up with more spontaneous lies for the ‘moving town’ scenario. There was however no significant main effect for condition $F$ (2, 147) = 1.62, $p = .202$, $\eta_p^2 = .02$ and no significant interaction between scenario type and condition $F$ (2, 147) = 2.58, $p = .128$, $\eta_p^2 = .03$. Descriptive statistics are presented in Table 4.5.

Table 4.5.

*Means and standard deviations for spontaneous lies in response to each scenario in each of the three experimental conditions*

| Scenario | Condition | | | Total |
|---|---|---|---|---|
| | Factual *M* (*SD*) | Upward *M* (*SD*) | Downward *M* (*SD*) | *M* (*SD*) |
| Moving town | 2.72 (1.34) | 2.82 (1.22) | 2.71 (1.63) | 2.75 (1.40) |
| Car crash | 1.86 (1.25) | 2.52 (1.34) | 2.41 (1.22) | 2.26 (1.26) |
| Totals | 2.30 (1.06) | 2.67 (1.04) | 2.56 (1.15) | - |

We also looked at the association between ratings of affect and number of lies for each scenario within each of the three conditions. We found no evidence for such a relationship (all p's > .218).

**Discussion Study 9**

The manipulation of affect was partially successful, and worked as intended, participants in the downward counterfactual condition felt more positive in comparison to participants in the upward condition as well as compared to those in the factual condition. There were no significant differences between those in the factual condition and those in the upward condition. A potential reason for this might be that even though participants in the factual condition were meant to focus solely on the outcome of the event they might not have been able to stifle the imagination of a better outcome, which often comes spontaneously (Roese & Olson, 1997). Thus, our manipulation might not have been strong enough to supress upward counterfactual thoughts in the factual condition. Similar to results in Study 8, participants felt worse in response to the 'car incident' scenario than the 'moving

town' scenario. As explained previously this might relate to the type of norm violation embedded within each scenario (Wilson & O'Gorman, 2003).

In terms of deception, overall, the counterfactual condition did not influence the number of lies participants spontaneously came up with. The pattern of results across conditions for each scenario appeared to be slightly different. A potential explanation for this difference is that lying to a neighbour is more acceptable than lying to the police, therefore for the baseline for deception in this scenario is already high and the manipulation might have not been strong enough to lift this above baseline. Indeed when looking at the number of lies in response to each scenario in the factual condition we found that whilst for the moving town scenario the number of lies ranged from 0 to 6, with most responses clustering around the 3 to 4 range, for the car crash scenario deceptive responses raged from 0 to 5 and the majority of responses clustered around the 1 to 2 range.

As opposed to study 8 there was a difference in spontaneous deception between the 'car incident' and 'moving town' scenarios with participants providing more lies in response to the 'moving town' scenario. However we did not find rating of affect to correlate to the number of deceptive responses provided by participants. Given that there was no difference between conditions in terms of deception this suggests that the motivation to lie was influenced by the risks of engaging in deception associated which differed between scenarios rather than the counterfactual nature of the thought participants engaged in. Overall, counterfactually derived affect did not influence subsequent deception.

This study has several limitations. One potential limitation to be addressed is the fact that participants were provided with five blank spaces to respond to

spontaneous deception questions. Although we reasoned that this was adequate in order to eliminate response interpretation bias, the existence of multiple spaces might have indeed primed participants to think about more alternatives than they would have otherwise. This could have masked the actual number of spontaneous lies one could have naturally thought about. Thus, due to this difference in method we might have assessed cued deception rather than spontaneous deception as in our previous studies (in Chapter 2). This interpretation is supported by the fact that the mean of deceptive responses in the current study is almost double that of deceptive responses from our previous studies.

The current study did improve on Study 8 by providing participants with the counterfactual thought and hence controlling for the type of available information within each condition. However the availability of information could have still been different across conditions. For example, counterfactual thoughts provided to participants in the upward condition could have easily been used to come up with a lie, whereas the thought provided to participants in the factual condition only focused on the outcome of the event and could only facilitate truth-telling. If the aim is to assess whether affect derived from counterfactual thoughts can facilitate deception beyond the availability of information than we should seek to investigate this effect when participants across conditions have the same type of information available to them. We will attempt to address this in Study 10. Our previous findings from Study 5 also pointed towards the potential impact of more specific moral emotions, such as guilt, on deception, and in Study 10 we sought to assess this.

## Study 10

In our previous studies we sought to investigate the influence of general affect in the relationship between counterfactuals and deception by manipulating the direction of counterfactual thoughts. Although we found some differences in the number of lies between counterfactual conditions, these could not be attributed to differences in affect derived from counterfactual thoughts. The current study was designed to assess specific moral emotions and to address some of the limitations in Study 9.

Apart from looking at emotions such as guilt, shame, regret and disappointment, we also aimed to expose participants to the same information so that we could limit the impact of the availability of information contained with counterfactual thoughts. Essentially, we wanted participants to have the same material available for constructing lies whilst assessing whether the affect which follows a counterfactual thought can influence whether they use this information or not. In order to achieve this we aimed to manipulate affect through causal attribution

The consequences of counterfactual thinking are mediated by attribution processes. Internal attributions lead to an increase in emotions such regret and whereas external attributions lead to an increase in disappointment (Zeelenberg et al., 1998). According to this research, thinking about how we could have changed things in our control so that a negative event could have been avoided should make us feel different comparison to thinking about how things outside of our control caused that same event. Guilt and shame have also been found to be related to internal attributions (Tracy & Robins, 2006).

Additionally, the methodology used to measure spontaneous deception in Study 8 and 9 might have masked genuine spontaneous generation. In order to make sure participants were not primed to come up with more lies that they would have spontaneously, we provided participants with only one space and asked them to write down any way in which they think they would respond to the neighbour/ police. This was similar to how we assessed spontaneous deception in Chapter 1.

If negative affect underlies the relationship between counterfactual thinking and lies then we expect we expect that deception will increase when counterfactual thoughts focus on internal attributions in comparison to when they focus on external attributions.

## Method

### Participants

A total of 214 took part in the study and completed all tasks. Just as in Study 8, because participants' accurate completion of sentence stems was an important part of the experimental manipulation, only participants who accurately completed the sentence stems were included. In total we had to exclude 64 (28.5%) of participants. Of the excluded participants 56 (26.1% of the entire study sample) were excluded from the low control condition. This was because they either completed sentence stems which included thoughts appropriate for the high control condition or because they did not provide complete sentences.

The remainder of 150 participants (104 females and 46 males) who were included in this study, ranged in age from 18 to 73 years ($M$ = 31.76, $SD$ =12.88),.

Each participant was monetarily rewarded at a rate of 0.12$ per study (the recommended rate by the recruitment platform).

## Procedure and materials

Participants were randomly assigned to one of three conditions, a baseline condition, a high control condition and a low control condition. In each condition participants were first instructed to read the two scenarios, i.e. the 'moving town scenario' and the 'car incident scenario' and vividly imagine that the events in the scenario happened to them. Scenarios were presented in a counterbalanced order.

**Counterfactual manipulation**. As before, each scenario lead to a negative event, however this time the scenarios were slightly modified to include precursors to the outcome that were either in the protagonist's control or not. Each scenario included two precursors which were within the protagonist's control and two which were not. For each scenario we made sure that undoing either the controllable or the uncontrollable events was sufficient to undo the overall negative outcome. In this way we wanted to make sure that focusing on either controllable or uncontrollable events could lead to the generation of plausible counterfactuals.

The moving town scenario read as follows:

*'**Recently the company you work for moved offices so you also had to move to a new city** where you don't know anyone. A few weeks after your move, your next-door neighbour invites you to a party. Most of the people who live on your road will be there. <u>You promise your neighbour you will attend</u>. Few days later, on your way to the party, the streets are wet and a **car drives too quickly past you and soaks you in water**. You have to go back home to change. When you get home you turn on the*

*TV and realize your favourite film is on, so you just decide to stay home and watch it instead of going to the party. Later that evening you realize that your move hasn't turned out anything like you had expected as you haven't made any real friends.'*

Whereas the car crash scenario was modified to be:

*'Friday, after a busy week at work you choose to go shopping. On your way back from the supermarket you* ***get diverted down a route which has narrow streets****. When you finally get closer to home,* ***your relative rings you on your mobile****. You decide to answer. As you try to pay attention to what they are saying you end up driving too close to the cars that are parked on the road. You hit someone's car and clip their wing mirror.'*

The events in bold were circumstances out of the protagonist's control whereas the one that are underlined were within their control.

Participants read each scenario at a time and answered questions relating to that scenario before moving on to the second scenario. After reading the scenario, depending on the condition they were assigned to participants were asked to read a sentence. For the high control condition participants were asked to simulate behaviour-focused counterfactuals. This was done by asking them to complete two sentence stems, starting with: If only I….the instructions read: 'Complete the following two sentences by indicating how aspects of your own choices, decisions or behaviour could have been different so that the event would have had a better ending.' Participants in the low control situation were asked to engage in situation-focused counterfactual thinking. This was done by asking them to complete to sentence stems, starting with: If only…. The instructions read: 'Complete the following sentences by indicating how aspects of the situation, thus things that you

could not influence yourself, could have been different so that the event would have had a better ending.' For the baseline condition participants did not have to complete any sentences and moved straight onto items which asked about emotional reactions.

**Measure of affect.** In order to assess how in control people felt in relation to the outcome of the scenario, given the counterfactuals they generated we asked four questions. Firstly, participants were asked 'How responsible do you feel for the event (not responsible to very responsible)?'. They were also asked 'To what extent did you cause the event?' and 'To what extend did external factors cause the event?'. Finally, we asked participants 'How much control do you feel you had over the outcome?' Participants responded to all these items on a 7 point Likert scale where 1 indicated low levels of the variable where 7 indicated high levels. We also assessed moral affect by asking participants to rate the amount of guilt, shame, regret and disappointment as a results of thinking either the behaviour-focused or situation-focused counterfactual thoughts. These were measured on a 7 point scale ranging from 1- not at all to 7 - very much.

**Measure of deception.** In order to assess spontaneous deception, participants were instructed to imagine either bumping into their neighbour who asks them why they didn't attend the party (moving town scenario) or being questioned by the police about their knowledge concerning the damaged car (car incident scenario). They read: *'Please use the space below to write down any way in which you think you would respond to the neighbour/ police.'* They were provided with one space in which they could enter free text.

## Results

The assumption of normality was violated for the outcome variables. Results were analysed using both parametric and non-parametric method. For the non-parametric analysis we used Generalized Estimating Equations (GEE). Pattern of results were similar for both analyses. For ease of interpretation we report the parametric analysis here.

### Perceived control results

To check whether the manipulation worked as intended we assessed whether the thoughts participants came up with in the various conditions made a difference in terms of how in control they felt over the outcome. We computed this score by reverse scoring responses to the question asking ‘To what extend did external factors cause the event?‘ and adding this score to the score from the other three items assessing perceived control. The resulting total control score could range from 4 to 28, with higher scores indicating a heightened perception of control. Responses were analysed using a mixed ANOVA analyses with scenario as the within-subjects factor (‘moving town’, ‘car crash’), and condition (baseline, high control, low control) as the between factor variable.

The analysis revealed a main effect for scenario $F$ (1, 147) = 118.05, $p < .001$, $\eta_p^2 = .45$, with participants feeling like they had more control over events in the car crash scenario ($M$ = 23.19, $SD$ = 3.91) than the moving town scenario ($M$ = 18.42, $SD$ = 5.25). There was no main effect of condition $F$ (2, 147) = .63, $p = .535$, $\eta_p^2 = .01$, and no significant interaction between scenario and condition $F$ (2, 147) = .89, $p = .414$, $\eta_p^2 = .01$.

**Affect results**

We calculated a total affect score by adding together the results from each of the four items which assessed moral emotions. This provided us with a total score which could range from 4 to 28, with higher scores indicating heightened negative moral affect. Descriptive statistics for moral affect score are provided in table 4.6. A mixed ANOVA released a main effect for scenario $F$ (1, 147) = 135.95, $p$ <.001, $\eta_p^2$ =.48, with participants feeling worse in regards to the 'car crash scenario' as compared to the 'moving town' scenario. There was no main effect of condition $F$ (2, 147) = 1.20, $p$ =.302, $\eta_p^2$ = .02, and no significant interaction between scenario and condition $F$ (2, 147) = 2.02, $p$ =.136, $\eta_p^2$ = .03.

Table 4.6.

*Means and standard deviations for moral affect for each individual scenario across all experimental conditions*

| | Condition | | | |
|---|---|---|---|---|
| | Baseline *M* (*SD*) | High Control *M* (*SD*) | Low Control *M* (*SD*) | Total *M* (*SD*) |
| Moving town | 16.52 (6.25) | 18.50 (5.18) | 18.67(5.25) | 17.87 (5.64) |
| Car crash | 23.44 (4.31) | 23.76 (4.67) | 23.31 (4.17) | 23.51 (4.36) |
| Total | 19.98 (4.22) | 21.13 (4.07) | 20.99 (3.91) | - |

We also performed the analysis with each type of emotion as a dependent variable. Overall the same pattern of results emerged other than for regret. In terms of regret, participants in the high control condition felt more regret than those in the

baseline condition but this was only the case for the moving town scenario. These results are at length presented in Appendix 4A.

**Spontaneous deception results**

To analyse differences in deception we performed a mixed factor ANOVA which included scenario as a within factor (moving town, car crash) and condition as a between factor (baseline, high control, low control). Descriptive statistics are presented in table 4.7.

There was no significant main effect for scenario $F$ (1, 147) = 2.38, $p$ =.125, $\eta_p^2$ =.02, or condition $F$ (2, 147) = .50, $p$ =.951, $\eta_p^2$ <.01. There also was no significant interaction between scenario and condition $F$ (2, 147) = .10, $p$ =.903, $\eta_p^2$ <.01. This pattern of results was similar both when introducing perceived control and affect as covariates.

Table 4.7.

*Means and standard deviations of lies for each individual scenario across all experimental conditions*

| | Condition | | | |
|---|---|---|---|---|
| | Baseline<br>*M* (*SD*) | High Control<br>*M* (*SD*) | Low Control<br>*M* (*SD*) | Total<br>*M* (*SD*) |
| Moving town | .94 (.65) | .92 (.85) | .92 (0.74) | .93 (.74) |
| Car crash | .76 (.87) | .76 (.77) | .83 (.86) | .78 (.83) |
| Total | .85 (.53) | .84 (.61) | .88 (.54) | - |

## Discussion Study 10

In the current study we aimed to manipulate moral affect through counterfactual thinking focused on different causal attribution. The manipulation was not successful, as participants did not engage in causal attribution differently across experimental conditions. There were also no differences between conditions for any of the moral emotions apart from regret. The consideration of undoing either controllable or uncontrollable events made participants feel more regretful in comparison to those that did not consider any counterfactual thoughts, however this was only the case for the 'moving town' scenario and not for the 'car incident' scenario. This might be because moral negative affect for the car incident scenario was already near ceiling in the baseline condition and thus the manipulation might not have been powerful enough to enhance this even further. We observed no difference in terms of regret between the high and low control conditions, which goes against research suggesting that regret would be more strongly associated with behavioural focused counterfactuals rather than situation focused counterfactuals (Zeelenberg et al., 1998). However, this is potentially because ratings of regret were most likely attributed to counterfactual thinking about the outcome rather than attribution to particular causes. That is, even though people were asked to focus on different elements that could have been undone they could have still only focused on the fact that things could have been better when answering the affect questions.

Similar to previous studies in this chapter participants felt more negative in response to the 'car incident' scenario in comparison to the 'moving town' scenario. These results are is in line with evidence which suggests that actions of

commission are more abnormal and possess a greater potential for evoking stronger emotional reactions (Kahneman & Tversky, 1982a).

In terms of deception, there were no differences between conditions, or scenarios. Therefore, this study could not find a link between counterfactually derived affect and deception. Even when we observed differences in regret, this did not appear to influence subsequent deception. The fact that we did not find any differences in deception between the baseline condition and the other two counterfactual conditions could be interpreted to be at odds with our previous conclusion about the effect of availability of information on deception. However, this could be explained by the fact that even though participants allocated to the baseline condition did not have to generate counterfactual thoughts in response to the scenario they still had to respond to manipulation check questions. Answering questions relating to perceived control might have required the consideration of how the outcome could have been different and hence alternatives might have been as available to participants in the baseline condition as they were to participants in the other conditions.

However, given that the manipulation did not work as intended limitations need to be addressed before a stronger conclusion can be drawn. Firstly the way in which emotions were assessed might have limited a differentiation between the different emotional reactions. Participants were asked to rate different moral emotions at the same time, and therefore it could be that these ratings influenced each other. One way to address this would be to assess the rating of affect between-subjects rather than within. Additionally, the manipulation might have not been strong enough for the car crash scenario as scores in the baseline condition, that is

the condition in which participants were not required to provide a counterfactual thought were already close to the highest rating possible.

## General discussion

The aim of the current chapter was to assess whether the negative affect derived from counterfactual thoughts can motivate subsequent deception. Evidence suggests that different emotional states follow such counterfactual thoughts depending on the factors such as direction, control and degree of comparison. The desire to deceive can also be amplified by such emotions

Our results were mixed as our manipulation of affect was not always successful. Overall, we found that when differences in affect were observed they did not translate into a difference in deception. Additionally, when differences in deception were observed these did not mirror differences in affect. Thus, we can tentatively conclude the way in which counterfactual thoughts make us feel in relation to a past outcome does not influence our tendency to lie about that outcome.

We have previously mentioned that counterfactual thought could be useful in deceptive communication because they make available information about the way in which events could have been different. This remains a plausible explanation given that we observed differences in deception between participants in the factual condition and those in the downward condition (Study 8), which could not be accounted for by differences in affect. Similarly, in Study 9, there were more deceptive responses in the counterfactual conditions as opposed to the factual condition, although this was mainly the case for the ‘car incident’ scenario. It is true that there were no differences between conditions in Study 10, however this could have been due to the fact that even in the baseline condition participants had to

engage in counterfactual thinking in order to respond to the question relating to causal attributions.

The current chapter also emphasised how contextual differences between scenarios can impact affective reactions to the scenario. These differences did not have a reliable effect on subsequent deception. There were differences in deception between the ‘car crash’ scenario and the ‘moving town’ scenario in Study 9, but this was no related to ratings of affect. This was more likely due to the level of risk associated with deception which differed between scenarios. Investigating the mediating effect of such contextual factors on the counterfactual thinking and deception relationship should be investigated further.

Other questions remain to also be answered given that the manipulation of affect did not always work as intended. Future studies should further seek to investigate the way in which counterfactually derived affect impacts deception. A better way to manipulate affect would be through behavioural counterfactual tasks such as the one used in Study 7, Chapter 3. This would allow participants to actually experience genuine emotions rather than imagine experiencing them. Additionally, future studies could seek to examine affect between-subjects so that scores for the different emotions do not contaminate each other. Until such limitations can be addressed, we would like to provisionally conclude that our studies did not find counterfactually derived affect to be an important factor in the relationship between deception and counterfactual thinking as assessed in this current thesis.

# CHAPTER 5

## Is executive function an underlying mechanism for the relationship?

This thesis has uncovered several findings. Firstly, we have established a direct association between counterfactual thoughts and deception by showing that counterfactual availability influences people's tendency to lie (Chapter 2). Secondly, we have shown that people use counterfactual thoughts in order to deceive and that those skilled in generating counterfactuals are more successful when lying to others (Chapter 3).

We have previously argued that counterfactuals motivate lying by increasing the general availability of alternatives to past events. However other mechanisms also have the potential to underlie this relationship. In Chapter 4 we attempted to investigate whether counterfactual thinking influences deception due to counterfactually derived negative affect, but we found little evidence for such a link. In Chapter 2, we also attempted to investigate whether executive function played a role in the association between deception and counterfactual thinking but found little evidence to suggest this. A potential explanation for these later findings rests on the characteristics of the population used. Namely, the undergraduate student population used in Study 1 was very able in terms of their executive function skill and there was little variance among the sample's scores. As such, in this final study we attempt to further investigate whether executive play functions a role in the relationship between counterfactuals and deception. We aim to do this by comparing the link between counterfactuals and deception in Parkinson's disease patients and age matched controls, two populations known to differ in their executive function skill.

Establishing executive function as an underlying mechanism for the relationship of interest could help us understand whether this link is due to the way in which we process and represent information.

## Executive Function and Counterfactuals

The production of counterfactual thoughts is interlinked with executive function performance. In order to come up with an alternative to the way things are, an individual needs to refrain from responding with the way in which things what actually occurred, keep active in mind both the factual and nonfactual, and be able to switch between what happened and what could have happened. Such skills rely on inhibitory control, working memory and switching respectively (Miyake et al., 2000).

Support for this comes from studies with children which have shown that the development of counterfactual thinking can be closely mapped onto the development of these executive functions (Beck, Riggs, & Burns, 2011). Better performance on inhibitory control tasks has been found to predict increased performance on counterfactual tasks (Beck et al., 2009). Similarly, attentional switching has been shown to impact comparisons between the real and counterfactual worlds which allow for the experience of counterfactually derived emotions such as regret (Burns et al., 2012). Evidence concerning the relationship between working memory and counterfactuals is not clear cut (Beck et al., 2009; Burns et al., 2012), however, some evidence does suggest that increased performance on working memory tasks is associated with better performance in response to counterfactual questions (Drayton et al., 2011; Guajardo et al., 2009).

In adults, findings are not as direct but nonetheless point in the same direction. For example, Whitfield, Turley-Ames, and Miyake (1999, as cited in Guajardo et al., 2009), have shown that imagining alternatives by removing an event which took place relies on working memory capacity. Xu, Bègue, and Bushman (2012) have also shown that participants who had their executive resources depleted, were less likely to feel guilt, an emotion linked to comparison of hypothetical alternatives (Niedenthal et al., 1994). Furthermore, as we grow older and our executive function skills decrease, our tendency to generate counterfactual thoughts also decreases (Walsh et al., under review).

Imaging studies have shown that counterfactual thinking activates brain networks in the prefrontal cortex (PFC), a brain region linked to major executive function deficits for executive function skills (for a review see Van Hoeck et al., 2015). For example, Coricelli et al. (2005) show that the orbitofontal cortex is involved in the experience of regret.

Further support comes from studies involving patients with damage to the PFC such as patients with Huntington's disease and schizophrenia who show an impaired ability to generate counterfactual thoughts and utilize counterfactuals to derive inferences (Hooker et al., 2000; Solca et al., 2015). In Huntington's disease, counterfactual thinking impairment was associated with decreased performance on measures of inhibition and switching (Solca et al., 2015). The experience of regret, which is known to result from the consideration of counterfactual alternatives, has also been shown to be affected in patients with orbitofrontal cortical lesions (Camille et al., 2004). Interestingly though, Beldarrain et al. (2005) also showed counterfactual thinking impairment in patients with PFC lesions although this was

linked to spontaneous generations only and not cued ones. This potentially suggests that these patients have the ability to construct counterfactual thoughts when explicitly asked but tend not to do so.

## Executive Function in Deception

The cognitive view of deception posits that lying is more mentally taxing than truth telling (Zuckerman et al., 1981 but see Burgoon, 2015). Not only do participants subjectively report this (Vrij, Semin, & Bull, 1996), they are also slower and more error prone when lying, indicators which denote increased cognitive effort (Vrij & Mann, 2001; Walczyk et al., 2003). In a similar manner to counterfactual thinking, there are several lines of evidence which indicate that lying is linked to executive function performance. It is postulated that lying requires one to refrain from stating the truth which tends to be spontaneously activated. Additionally, in order to lie one needs to keep multiple pieces of information active in working memory such as the receiver's perspective as well as the information that could constitute a lie. Furthermore, one also has to skilfully switch between the different pieces of information (Gombos, 2006).

In support of this, research shows that liars who are better at convincing others perform better on verbal working memory tasks and present with more efficient inhibition skills (Alloway et al., 2015; Debey, Ridderinkhof, De Houwer, De Schryver, & Verschuere, 2015). Additionally, those who are practiced at switching between telling the truth and lying are less likely to make errors when responding deceptively (Van Bockstaele, Wilhelm, Meijer, Debey, & Verschuere, 2015). Studies which have directly manipulated executive functions similarly found that lying is sensitive to cognitive demands such as inhibiting the goal-irrelevant

truth response (Debey, Verschuere, & Crombez, 2012). Moreover, across the life-span the propensity to deceive follows the inverted U shape trajectory of inhibitory control ability, it peaks in adolescence and declines subsequently (Debey, De Schryver, et al., 2015).

The higher cognitive cost of lying is also reflected in higher activation of PFC areas known to correlate with performance on executive function tasks (Abe, 2009; Christ et al., 2009; Spence et al., 2004). For example, the PFC areas implicated in working memory have been found to activate when participants were instructed to be dishonest in recognizing a picture stimuli (Ito et al., 2011). PFC areas linked to switching and response inhibition also shows significant activation when participants chose to lie about personal relevant material (Spence, Kaylor-Hughes, Farrow, & Wilkinson, 2008).

Additional support comes from research on clinical populations and studies which have experimentally manipulated neural activity in the PFC with current stimulation. Yang et al. (2005) show that pathological liars (i.e. those that repeatedly and compulsively tell false stories) present with an increase in white matter in the PFC. Patients with PFC lesions show an impairment in their ability to infer deception (Stuss, Gallup, & Alexander, 2001). Priori et al. (2008) additionally found that the use of transcranial direct current stimulation (tDCS) to stimulate the dorsolateral PFC decreased the speed with which participants lied about visual stimuli when told to do so. Similarly, Karton and Bachmann (2011) found that stimulation of the brain area through repetitive transcranial magnetic stimulation resulted in a decrease in the propensity to lie. Conversely however, Mameli et al. (2010) found that stimulation of the dorsolateral PFC through tDCS actually

facilitated lies, in that participants lied quicker when cued to do so. This however was only the case for lies concerning general knowledge and not lies about personal knowledge. Such findings indicate that different types of lies are linked to distinctive neural mechanisms, and hence different executive functions (Abe et al., 2006; Ganis et al., 2003). Overall the widely held view is that deception relies on executive functions and that this might differ depending on the type of lie being assessed.

### Counterfactual thinking and deception in Parkinson's disease

As mentioned above, counterfactual thinking and deception have been shown to be impaired in various populations of patients with prefrontal lobe lesions. One particular population which has been found to be impaired in both the generation of alternatives to the past and in the ability to tell lie is Parkinson's disease (PD) patients.

PD is a slowly progressing neurodegenerative disorder primarily characterized by motor disturbances, namely slowness and poverty of movement. In addition, PD patients also present with non-motor disturbances (Jankovic, 2008). These impairments include hallucination, delusions and psychosis as well as cognitive impairment related mostly to executive function deficit. Patients with PD have been found to exhibit issues with internal control of attention, set shifting, inhibitory control and dual task performance (Dirnberger & Jahanshahi, 2013).

These executive function deficits have also been linked to the ability to produce counterfactual thoughts and deception in this patient population. McNamara et al. (2003) found that compared to controls PD participants had fewer thoughts about how things could have been different when considering a past

personal negative event. Furthermore, patients were also impaired in their ability use counterfactuals in order to draw inferences. This performance was correlated with impairment on tests of executive function, such as working memory and verbal fluency and ta decrease in social functioning.

In a separate study, PD patients were also found to be impaired in their ability to lie. Abe et al. (2009) found that in comparison to controls, PD participants lied less about certain visual stimuli when instructed to do so. Additionally, this deficit was linked to reduced brain activation in the dorsolateral PFC. Together these two studies suggest once more that counterfactual thinking and deception are linked to executive function. They also underscore a potential hypothesis, namely that a lack of counterfactual thinking ability is associated with impairment in deception and that executive function abilities may partly explain this relationship.

However, before investigating such a hypothesis directly, several limitations concerning the existing research need to be taken into account. Firstly, in the McNamara et al. (2003) study, PD participants and controls differed significantly on measures of global cognitive function that go beyond executive function deficits. More specifically, PD participants' scores were significantly lower than those of controls on the Mini-Mental State Examination (MMSE), a measure of basic cognitive ability (Folstein, Folstein, & McHugh, 1975). This is relevant given that the PD group score on average of 26.78 out of a maximum total possible score of 30, and that the diagnostic criteria for mild cognitive impairment (MCI) in this population is a score of 26 (Litvan et al., 2011). This is additionally notable because the task assessing counterfactual thinking relied on episodic memory which has been found to be impaired in patients with MCI (Pfeiffer, Løkkegaard, Zoetmulder,

Friberg, & Werdelin, 2014). Questions also remain about whether PD patients have an impaired in the generation of spontaneous counterfactual thoughts, given that in their study McNamara et al. (2003) only assessed participants' thoughts about how things could have been different when directly asked to do so. These questions are more pertinent given that other researchers have not found such impairments in PD participants when using vignette based tasks (Eddy, Beck, Mitchell, Praamstra, & Pall, 2013). Also to be considered that PD patients can excel on test which require the imagination of alternative possibilities when treated with dopaminergic drugs (Faust-Socher, Kenett, Cohen, Hassin‐Baer, & Inzelberg, 2014).

Secondly, in the study by Abe et al. (2009) deception was assessed only in regards to questions which required only simple yes /no responses, thus leaving the question of whether this patient population is also impaired in the production of lies involving a complex narrative. Additionally, participants were externally cued for whether to lie or tell the truth and thus there is no evidence for whether PD patients differed in their spontaneous propensity to deceive. Furthermore, participants' responses had no tangible consequence which makes the process of lying assessed in this study fairly different from real-world deception.

Overall both deception and counterfactual thinking have been found to be impaired in PD. However, questions concerning the nature and extent of such impairments remain.

## Study 11

Given the evidence reviewed above, the current study was primarily designed to assess whether impairment in counterfactual thinking is mirrored by impairment

in deception. We aimed to do so by examining the performance of PD participants in comparison to healthy controls in order to further assess whether executive functions was the basis for this relationship. In doing so, the current study also sought to replicate findings suggesting that PD patients are impaired in their generation of counterfactuals and lies. Based on our previous findings we hypothesise a positive relationship between deception and counterfactual thinking.

## Method

### Participants

Eighteen participants with PD (9 female and 9 male) were recruited through the South-West branch of the Parkinson's UK Society. All patients reported to have been diagnosed with PD by a neurologist. Mean disease duration, defined as time since diagnosis, was 7.67 years. All patients were under dopaminergic drug treatment at the time of testing. Exclusion criteria were: history of substance abuse and concurrent neurological or psychiatric illness.

Eighteen gender matched controls (9 female and 9 male) were recruited though Parkinson's UK Society meetings (these were careers of patients with PD) and the University of the Third Age in Plymouth. Participants were matched for age (PD: Range = 51 to 83, $M$ = 68.22, $SD$ = 8.00; Controls: Range = 53 to 83, $M$ = 67.44, $SD$ = 7.56; $t(34) = .30$, $p = .766$) and gender. There were no significant differences between groups in terms of education $\chi^2$ (5) = 2.53, $p = .771$. None of the participants were not rewarded for their participation.

### Measures and procedure

All participants completed a battery of tests which included measures

of counterfactual thinking and deception as well as measures of executive function, disease severity and general cognitive skill.

**Counterfactual thinking measures.** Both spontaneous and cued generations of counterfactual thoughts were assessed through the use of the 'moving town' and 'car incident' scenarios utilized in Study1. Participants had to read the scenarios and imagine the events actually happened to them. After reading the scenarios participants were asked to reflect back on events and imagine they were writing a diary page. This was done in order to assess spontaneous counterfactual thoughts. Instead of responding in writing, participants responded verbally and responses were tape recorded. We recorded verbal responses instead of written ones because PD participants can exhibit impaired handwriting kinematics (Phillips, Stelmach, & Teasdale, 1991). The order in which the scenarios were presented was counterbalanced. To measure cued generations participants were directly instructed to imagine as many ways as possible that events could have been different. Participants could re-read the scenarios during this task. Counterfactuals were coded using the same coding scheme as in Study 1 (Appendix 2B). Two independent raters coded the spontaneous generations and inter-rater agreement was high (r=.94). Raters were not blind to hypothesis but were blind to experimental condition. All discrepancies were resolved by discussion.

We also included the *Counterfactual Inference Test (CIT)* developed by Hooker et al. (2000) in order to test individual's tendency to use available counterfactuals. The *CIT* includes four forced choice questions each with three response options. Each item represents events experienced by two individuals for

which circumstances differ such that one has a more readily available counterfactual alternative than the other. We recorded the total number of items for which participants gave the answers which relied on counterfactual inferences.

**Deception measures.** Spontaneous and cued lies were assessed using the same two scenarios used in Study 1 (Appendix 2A). For spontaneous generation, participants were asked both a direct and indirect question. The indirect question asked what they would say to the neighbour or the police following the wrongdoing presented in the scenarios. The direct question asked participants to state what they would say if they were specifically questioned about why they didn't attend the party or what they knew about the car that was damaged. For cued generation participants were asked to list as many lies as they could think of for each scenario. Questions asked are presented in Appendix 2B. Lies were coded according to the previously used definition of deception by Serota et al. (2010). Each unique lie was given a score of 1.

Participants were also asked to report the number of lies they told in the last 24 hours to either family, friends, acquaintances and total strangers either face to face or in situations mediated by technology (internet or phone). We recorded the total number of lies across all of these situations (Serota et al., 2010).

**Executive function measures.** Phonemic verbal fluency was tested using the *FAS test (Strauss et al., 2006)*. Participants were asked to produce as many words beginning with the letter F and then with the letter A, and finally with the letter S within a time limit of 1 minute per letter. Participants were asked to avoid producing names of people or places and repetitions of words. Letter fluency was chosen as opposed to category fluency because this task has been shown to rely

more on executive function than verbal ability (Pettit, McCarthy, Davenport, & Abrahams, 2013). The score recorded was the total number of unique correct words across all three trials

The *Stroop test* was used in order to measure inhibitory control (Golden & Freshwater, 2002; Stroop, 1935). The test consisted of two separate trials. In the first trial, participants had to read out-loud as quickly as possible a series of words printed in congruent colours (C trial). In the second trial, participants were given a different set of colored words and asked to read the colour the word was printed in (C-W trial). The C-W trial is also referred to as the interference trial as it requires participants to consciously override the intrusive drive to read the word itself and name the colour of the ink instead. The time taken to complete each trial is recorded in seconds and milliseconds. In the McNamara et al. (2003) study on the C-W trial score alone was used in order to assess inhibitory control. However, such a score has been shown to be less informative in populations with impairment in processing speed such as PD patients. Low interference scores in these groups may be indicative of general cognitive slowing rather than response inhibition per se (Denney & Lynch, 2009). Therefore, in order to control for confounds associated with processing speed, we used the ratio score. This score divides the time spend on the C-W by the time on spend on the C trial. Because it can account for processing speed and because it can be easily calculated, the ratio score has been recommended as the most appropriate for research (Sisco et al., 2016). Higher Stroop ratio scores indicate greater impairment. The computerized version of the Stroop has been found to diminish interference effects, and so we used the original paper version (Penner et al., 2012).

We also used the *Trail Making Test (TMT)* (Reitan, 1958) in order to measure cognitive flexibility. The test involves two parts. In part A participants are required to draw a line between numbered circles in numerical order (i.e. 1, 2, 3, etc.) as quickly as possible. In part B, participants had to draw lines between circles with numbers and letters in alternating order (i.e. 1-A-2-B-3-C, etc.). The time in seconds taken to complete each part represents the outcome variable. TMT is multidimensional in nature and involves motor speed, visual scanning and sequencing as well as cognitive flexibility. In an attempt to isolate the cognitive flexibility element, the current study calculated TMT score as a ratio between TMT-A and TMT-B (Golden et al., 1981). Higher TMT ratio scores indicate greater impairment in cognitive flexibility.

Verbal working memory was tested using the *Digit Span test* from the Wechsler Memory Scale-Revised (WMS-R) (Wechsler, 1997). The test requires participants to repeat a sequence of numbers presented to them. For the Forward subpart participants need to repeat the sequence in the same order in which it was presented to them. For the Backwards subpart the sequence needs to be presented in reverse. Total score was obtained by summing the scores of the two subparts. The Backwards subpart has also been validated as a diagnostic tool for determining cognitive impairment in PD (Warden, Hwang, Marshall, Fenesy, & Poston, 2016).

**General cognitive measure.** The *MMSE* (Folstein et al., 1975) was used in order to assess global cognitive function. The MMSE is a short test which assessed registration, attention and calculation, recall, language, ability to follow simple commands and orientation. The highest possible score is 30 and in

Parkinson's disease a score of 26 is considered a cut of point useful in indicating MCI (Litvan et al., 2011).

## Results

### Executive and cognitive function results

Table 5.1 lists the executive function and general cognitive ability results alongside the statistical differences between PD patients and controls. PD patients did not differ from controls on any neuropsychological measures other than the TMT test. Although there were significant differences in terms of the C-W interference Stroop trial, there was no difference when using the ratio score which takes into account processing speed.

### Counterfactual thinking results

We first checked whether participants differed in the way they engaged with the scenarios by comparing the number of words participants generated for the diary tasks. Results revealed that across the two scenarios, both groups responded similarly in terms of effort, $M$ = 314.00, $SD$ = 176.20 for PD and $M$ = 365.56, $SD$ = 179.62 for controls, $t$ (34) = -.92, $p$ = .366.

Although there was a trend towards PD participants coming up with fewer counterfactual thoughts in comparison to controls ($M$ = .72, $SD$ = 1.07) (range 0 to 4)[6] vs ($M$ = 1.56, $SD$ = 1.72) (range 0 to 6) respectively, this difference was not statistically significant $U$ = 113.50, $p$ = .103. For the numbers of cued generations,

[6] Outlier analyses identified an outlier in the spontaneous counterfactual scores for the 'car crash scenario'. Considering the type of population and the sample size we decided not to remove this outlier. Removal of this outlier would have resulted in a marginally significant difference between groups in terms of spontaneous generations (p = .051).

PD participants (*M* = 5.28, *SD* = 3.43) did not differ significantly from controls (*M* = 5.82, *SD* = 2.09) *U* = 123.50, *p* = .217.

Table 5.1

*Neuropsychological test results and statistical differences between PD patients and controls*

| | PD patients | Controls | P value |
|---|---|---|---|
| | *M* (*SD*) | *M* (*SD*) | *M* (*SD*) |
| MMSE | 29.17 (0.79) | 29.22 (0.94) | .849 |
| FAS | 45.94 (17.11) | 45.56 (12.11) | .938 |
| Digit Span | | | |
| Forward | 11.06 (2.18) | 11.17 (2.20) | .880 |
| Backward | 7.06 (2.41) | 7.94 (2.55) | .291 |
| Trail Making | | | |
| TMT A | 41.60 (17.81) | 31.32 (10.05) | .040 |
| TMT B | 109.76 (65.98) | 54.37 (13.77) | .003 |
| Part B/ Part A | 2.61(0.92) | 1.82 (0.52) | .004 |
| Stroop | | | |
| Stroop C | 66.48 (21.46) | 55.30 (9.23) | .050 |
| Stroop CW | 182.25 (64.76) | 133.87 (23.24) | .005 |
| CW/C | 2.71 (0.66) | 2.45 (0.48) | .222 |

*Note:* one PD participant could not complete the CW Stroop task due to colour blindness

In response to the CIT items, responses of PD participants (*M* = 1.72, *SD* = 1.23) were not significantly different in comparison to responses from controls (*M* = 1.67, *SD* = 1.28) *U* = 158.50, p = .909.

Correlations were performed to examine the relationship between performance on neuropsychological tasks and counterfactual thinking tasks in the Parkinson's population. Results are presented in table. 5.2. Only the score on the

Digit Span test was positively correlated with the generation of spontaneous counterfactual thoughts.

Table 5.2

*Correlation between neuropsychological test scores and measures of counterfactual thinking for the PD population.*

| | Spontaneous CFT | Cued CFT | CIT score |
|---|---|---|---|
| MMSE | .19 | -.25 | -.10 |
| FAS | -.31 | .39 | .29 |
| Stroop ratio score | .05 | .40 | -.05 |
| TMT ratio score | -.22 | -.07 | -.27 |
| Digit Span Total | .52* | -.23 | -.40 |

* $p < .05$

We also tested the association between counterfactual measures and measures of executive functions in the overall sample whilst controlling for group membership. We found a similar pattern of results across all correlations other than the correlation between Digit Span total score and spontaneous generations of counterfactuals for which the association was not significant ( $r = .14$, $p = .420$).

**Deception results**

PD participants ($M = 1.89$, $SD = 1.07$) and controls ($M = 1.83$, $SD = 1.04$), did not differ significantly in the number of spontaneous generations of lies $U = 160.50$, $p = .961$. There were also no significant differences between participant groups ($M = 4.22$ , $SD = 2.07$) vs ($M = 4.50$ , $SD = 1.47$) for PD and controls respectively, in terms of cued generations $U = 140.00$, $p = .478$.

In terms of lies told in the past 24h, there was a trend indicating that PD participants reported telling fewer lies ($M = .28$ , $SD = .67$ ) in comparison to

controls, ($M$ = .67 , $SD$ = .84) however this difference not statistically significant, $U$ = 119.00, $p$ = .094.

As revealed in table 5.3, there were no significant associations between any of neuropsychological test scores and measures of deception, apart for the TMT ratio score which positively correlated with the number of spontaneous lies participants came up with.

Table 5.3

*Correlation between neuropsychological test scores and measures of deception for the PD population.*

| | Spontaneous Lies | Cued Lies | Lies in past 24h |
|---|---|---|---|
| MMSE | -.30 | .21 | .31 |
| FAS | .29 | .01 | .36 |
| Stroop ratio score | .17 | .15 | .14 |
| TMT ratio score[a] | .48* | -.09 | -.18 |
| Digit Span Total | .11 | .11 | .07 |

[a] larger TMT ratio scores indicate greater impairment
* p < .05

We also tested the association between deception scores and executive functions scores in the overall sample whilst taking into account group membership. We found a similar pattern of results across all correlations other than the correlation between TMT ratio score and spontaneous generations lies for which the associations was not significant ( $r$ = .13, $p$ = .463).

**Association between deception and counterfactuals in PD**

Correlations were conducted to examine relationship between counterfactual thinking and deception in PD patients. Results revealed that spontaneous generations of counterfactuals and spontaneous generations of lies were negatively associated ($r$ = -.47, $p$ = .049). The trend for the association between cued generations of counterfactuals and deception was positive but this association was not statistically significant ($r$ = .41, $p$ = .096). Correlations between all measures of counterfactual thinking and deception are presented in table 5.4.

Table 5.4

*Association between counterfactual thinking measures and deception measures*

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 1.Spontaneous CFT | | | | | |
| 2. Cued CFT | -.43 | | | | |
| 3. CIT score | -.16 | .24 | | | |
| 4. Spontaneous Lies | -.47* | .01 | -.25 | | |
| 5. Cued lies | -.18 | .41 | .15 | .04 | |
| 6. No lies in past 24h | -.07 | .15 | .13 | -.07 | .32 |

*p < .05

When controlling for number of words spoken across the two diary tasks, the association between spontaneous counterfactual thinking and deception was only marginally significant ($r$= -.31, $p$ = .068), whilst the correlation between cued generations was not statistically significant ($r$ = .22, $p$ = .214)

We also tested the association between responses to counterfactual thinking measures and deception measures in the overall sample whilst taking into account group membership. We found a similar pattern of results across all correlations other

than the correlation between spontaneous generations of counterfactuals and cued counterfactuals generations which indicated a weaker association ($r$ = .01, $p$ = .995). The correlations between spontaneous counterfactuals and spontaneous lies pointed towards a negative association again although this time results were only marginally significant ($r$ = -.33, $p$ = .052). Furthermore, the association between spontaneous generations of lies and the number of lies told in the past 24 hours became significant ($r$ = -.37, $p$ = .029).

## Discussion Study 11

In the current chapter we aimed to assess whether impairment in counterfactual thinking was mirrored by an impairment in deception and whether such a relationship could be underlined by executive function skills. Our findings suggest that PD participants were not impaired in either counterfactual thinking or lying when compared to age-matched controls. Furthermore, we found little connection between such skills and executive function. Similar to our previous studies, there was an association between counterfactual thinking and deception however in contrast to our previous studies this was negative. Several potential explanations exist for why findings from this study differed from previous findings in the wider literature and from studies in the current thesis, we next seek to present these in turn.

Firstly, our findings do not match those from the McNamara et al. (2003) study as we didn't find PD participants to be impaired in the cued generations of counterfactuals or in the way they responded to the CIT test. Furthermore, even though a difference could be observed in terms of spontaneous generations, this did not reach statistical significance. Limited statistical power because of the modest

sample size in the present study (*N* = 36) may have played a role in limiting the significance of some of the statistical comparisons conducted. A post-hoc power analysis revealed that on the basis of the mean, between- groups comparison effect size observed in the study by McNamara et al. (2003) ( d = .84), a total sample of n of 50 would be needed to obtain a statistical power at the recommended .80 level (Cohen, 1988)[7].

Further explanations could rest on the characteristics of the sample and the methodology used. Our patient sample was arguably less impaired. Even though they had impairment in cognitive flexibility as well as inhibition (when assessed in a similar manner to the original study), our PD participants were not impaired in verbal fluency or general cognitive ability in comparison to controls. This might imply that our patient population was not impaired sufficiently in order to display counterfactual thinking deficits. Conversely, it could also be argued that the differences observed in the study by McNamara et al. (2003) were due to a generalized cognitive deficit rather than a specific deficit in executive function.

The methodology used also needs to be taken into account. In order to address concerns about differences in the ability to retrieve and generate a detailed account of a negative personal event, our study used hypothetical scenarios. We thus assessed counterfactual thoughts related to non-personal events rather than episodic ones as in the original study. Studies which similarly found no counterfactual impairment in PD also used vignettes (Eddy et al., 2013). This could be relevant given that patients with prefrontal lobe damage have been found to be impaired in the generation of self-relevant counterfactual thoughts but not impaired in the use of

[7] Our study had a similar statistical power compared to the study by McNamara et al., 2013. Namely, our study had a power of 66% whereas the previous study had a statistical power of 70%.

counterfactuals in a non-personal task (Beldarrain et al., 2005). Similarly, evidence suggests that different neural mechanisms are involved in the simulation of personal and impersonal counterfactual thoughts (De Brigard, Spreng, Mitchell, & Schacter, 2015). This is also potentially why overall, we found executive function scores were not to be associated with measures of counterfactual thought as hypothesised. Although we observed an association between working memory and spontaneous counterfactuals this might have been due to demands related to keeping in mind both scenarios in order to answer adequately. It is worth emphasizing however that our study also failed to find differences between responses to the CIT which was also used in the original study. Therefore it seems more likely that our contrasting findings are due to differences in general cognitive impairment than differences in the methodology used. The differences in methodology though do highlight that our results do not necessarily show that PD participants are able to think counterfactually, but rather that they are able to generate such thoughts when considering other people's experiences.

An ability to engage in counterfactual thinking was mirrored by an ability to deceive, our PD participants could come up with lies just as much as their age-matched controls. The ways in which the two studies differed could explain the discrepancy. The paradigm for assessing deception used in this study was advantageous because it allowed participants to choose whether to lie or not, whereas previous studies only instructed people when to do so. Additionally, the scenarios included a motivation to lie and participants were required to generate more complex narrative lies rather than simple yes/no answers. Despite these advantages it is likely that our deception task did not pose as much on demand executive function skill as compared to tasks used in previous studies. Participants

did not have to answer under time pressure and even though they responded out loud they were not directly deceiving the experimenter and thus did not have to monitor their body language or to make sure the lie was cohesive. This implies that there might have been little reliance on inhibitory control or working memory, a hypothesis corroborated by the fact deception responses were not associated with any such executive function scores. There was a correlation between TMT ratio score and spontaneous lies indicating that those who were less cognitively flexible told more lies. This contradicts previous findings which argue that cognitive flexibility should be positively linked with the ability to lie (Abe et al., 2009). However, if we consider that our spontaneous deception task could have been influenced by social demands (a limitation which we will discuss at length further on in this chapter), those less cognitive flexibility could have lied more because they could not accommodate to the social context. Having said that, we need to interpret such results with caution given the small sample size and the fact that even measures meant to assess similar variables, such as the different deception measures, did not correlate with each other either.

Interestingly in comparison to PD participants, controls did report telling more than double the amount of lies in the past 24 hours. Even though this difference was not statistically significant this does potentially indicate that the method for assessing deception is an important element to be considered. Overall our results show that PD participants are just as able as healthy age-matched controls to create complex narrative lies when assessed with scenario based measures.

Lastly the discrepancy with our own findings needs to be explained. Unlike in our previous studies we found a negative association between spontaneous lies

and the spontaneous generation of counterfactuals. Those who generated more counterfactual thoughts told fewer lies. The current study differed from our previous studies in two significant ways. Firstly the participants this this study were older than those in Study 1, and secondly participants responded to counterfactual and deception questions out-loud rather than in writing. These discrepancies could explain the difference in results.

Given the difference in participants' age it could be hypothesised that the relationship between counterfactuals and deception follows an inverted U shape relative to age. This interpretation could be in line with evidence suggesting the way in which we interpret social rules in different types of situations differs with age (Blanchard-Fields, Mienaltowski, & Seay, 2007). However, this hypothesis is provided tentatively, and future studies should seek to investigate the way in which factors associated with age have the potential to modulate the relationship between counterfactual thinking and deception.

An alternative explanation could rest on the way in which we recorded responses. To assess spontaneous deception participants were asked how they would act in situations which emphasized ethical standards. It is therefore likely that asking participants to respond out loud increased self-awareness as well as their desire to appear moral in front of the experimenter. Previous research has shown that situational factors can influence participants' desire to appear moral even though they might not intrinsically have a preference for honesty (Hao & Houser, 2011). Given that counterfactual thinking can facilitate social functioning (Epstude & Roese, 2008), it could be that the desire to appear moral in front of the experimenter manifested more strongly in those able to imagine alternative possibilities, which is

why we observed a negative relationship between counterfactuals and lying. This interpretation is supported by the fact that the association between cued generations followed the opposite pattern. In the cued generation task, participants were asked to generate as many lies as they could think of, this potentially removed the social desirability bias because it does not emphasise participants' moral fabric but rather their ability generate hypothetical lies. The different characteristics of the two types of task could also explain why cued and spontaneous generations had the opposite pattern of results in relation to executive function measures.

Across the interpretation of all the results in this chapter consideration needs to be given to the fact our participants were tested whilst on dopamine treatment. Dopamine has been shown to be beneficial to cognitive performance although such benefits are task-related, (Gotham, Brown, & Marsden, 1988; Lewis, Slabosz, Robbins, Barker, & Owen, 2005). For example, PD participants tend to be impaired on verbal fluency in comparison to controls only when off dopamine treatment. This is potentially why we also observed inconsistent results across the executive function tasks. Others have also shown that PD participants treated with dopaminergic drugs can demonstrate enhanced verbal creativity in comparison to healthy controls (Faust-Socher et al., 2014). Additionally, our participants were recruited through community settings, which might also imply they were more socially active and more cognitively skilled with perhaps more opportunities to practice deceptive skills.

In conclusion, the results of this chapter suggest that counterfactuals and deception are associated but further suggest that the direction of the relationship could be influenced by age or situational characteristics such as social desirability. We did not find this relationship be explained by executive function skill. Future

studies could seek and to address this question by using a behavioural paradigm such as the one utilized in Study 7.

# CHAPTER 6

## General Discussion

The principal aim of this research was to directly investigate the link between counterfactual thinking and deception. A secondary aim was to understand the potential mechanisms underlying this relationship. In the current chapter we will review all experimental findings in this thesis and interpret them in relation to existing research. We will also aim to highlight how this research contributes to our understanding of counterfactual thinking and deception. Furthermore, we will discuss potential limitations and suggest avenues for future investigation.

### Summary of findings

In Chapter 1 we reviewed separate lines of research which suggest similarities between counterfactual thinking and deception, for example both involve mentally altering past events (Byrne, 2016; Debey et al., 2014), and both have been reported to rely on similar mechanisms (Drayton et al., 2011; Gombos, 2006). Despite this, only a few studies assessed counterfactual thinking and lies concurrently. Shalvi et al. (2011) for example found that individuals who observe desired counterfactuals tend to lie more about their performance for monetary gain. These findings have more recently been replicated by Bassarak et al. (2017) who have also shown this effect persists even when the desired counterfactual is produced by someone else. Whilst the findings by Bassarak et al. (2017) imply that the mere imagination of counterfactuals can motivate lying, thus far there have been no empirical studies to assess whether the mental simulations of counterfactuals can

also affect the generation of lies. Furthermore, the previous studies have only inferred deception and counterfactual thinking through population-based performance which does not allow for analyses of data at an individual level. In light of this, the studies in this thesis are the first to investigate the relationship between the mental simulation of counterfactual alternatives and deception with measures which allow for the observation of counterfactual generations and lies at an individual level. Additionally, the current thesis also includes the first studies which aim to directly examine whether counterfactual thinking and deception share similar processes.

**The link between counterfactuals and lies**

In Study 1 we assessed this relationship using an individual differences approach and directly evaluated the generation of counterfactual thoughts and lies using scenario-based methods. We found that individuals' tendency and ability to generate counterfactuals is associated with a similar tendency and ability to produce lies. This evidence is in line with findings showing that that creative thinkers and those with higher cognitive abilities are more likely to deceive (Gino & Ariely, 2012; Sarzyńska et al., 2017).

In Study 6, using a similar individual difference approach we assessed the relationship by implementing a behavioural deception paradigm which allowed our participants to actively lie to others. The DeceIT paradigm did not adequately motivate participants to choose to lie when given the option, as it was more advantageous for them to choose to tell the truth in order to win a prize. This meant that results from Study 6 did not reveal a relationship between counterfactuals and the choice to lie. However, we did observe that those with a tendency to come up

with counterfactuals were also more likely to generate lies that were less detectable, suggesting that those with an increased tendency to think counterfactual thoughts have a greater deceptive ability. This evidence is in line with research suggesting that reasoning serves the function of persuading others and that counterfactuals can enhance the persuasive impact of a narrative (Mercier & Sperber, 2011; Tal-Or, Boninger, Poran & Gleicher, 2004). Taken together results from studies 1 and 6 are the first to show a direct link between deception and the mental simulation of counterfactuals. These results also are the first to emphasize that counterfactual thinking is an important process in the generation of lies.

The results based on individual differences were further supported by our experimental work which focused on manipulating the availability of counterfactual and investigating the impact on subsequent deception. Study 3 showed that in situations in which counterfactual information is highly available, participants tend to generate a greater numbers of lies. Study 7, which used a behavioural measure of deception, provided further support by showing that in these situations individuals do use these lies in order to mislead another. In particular, we showed that when missing a desired outcome individuals tend to lie more when missing the outcome by a narrow margin rather than a wider margin. These findings suggest that our moral compass can be swayed by the counterfactual nature of a situation and therefore provide more information as to how our intrinsic honesty can be compromised. Whilst previous research has already shown that people tend to lie more in situations where desired outcomes are narrowly missed (Hilbig & Hessler, 2013; Shalvi et al., 2011), our studies are the first to directly show that counterfactual thinking is the cognitive mechanism that underpins such findings.

We also aimed to enrich evidence about the relationship between counterfactuals and deception by assessing whether counterfactual thinking can influence the perception of dishonest behaviour. Given that counterfactuals can influence the way we judge other's intentions and traits (Knobe, 2010; Ndubuisi & Byrne, 2013), Studies 4 and 5 analysed whether people judged others as more likely to lie following situations where counterfactuals are more salient. Results from Study 4 showed that protagonists were judged as more likely to be deceptive when counterfactuals were more readily available. Study 5 replicated these results and clarified such judgements were indeed based on the counterfactual nature of the situation. Results in Study 5 also link to those in Study 6 as participants sometimes reasoned that lying following a situation in which counterfactuals are readily available, would be easier to tell and more credible. Overall, results from Studies 4 and 5 also agree with findings by Miller et al. (2005) and extend them by showing that we tend to judge others as more dishonest when we consider that they are more likely to imagine counterfactual alternatives.

**Underlying mechanisms**

The secondary aim of the current thesis was to investigate the underlying mechanisms of this relationship. In our studies, the content of participants' counterfactual thoughts did not necessarily match the content of their subsequent lies. This led us to hypothesise that counterfactuals might in part influence lies through a content-neutral pathway, that is indirectly through mind-sets and motivations (Epstude & Roese, 2008). Namely, we argue that thinking about how things could have been different increases the general availability of alternatives to past events which in turn aids the generation of lies.

One might argue that creativity could provide a possible explanation for why individuals who tend to think counterfactually also tend to lie. Creativity has indeed been associated with deception (Gino & Ariely, 2012) although some have questioned this link (Niepel et al., 2015). Similarly, creativity is also associated with certain types of counterfactual thoughts (Markman et al., 2007). However, many counterfactual thoughts and lies are not creative (e.g., those that remove an event from a scenario) (Gaut, 2003), and therefore this is unlikely to be the full explanation for the results.

Another possible interpretation which could also be accounted by the content neutral pathway is that counterfactuals might influence dishonesty through negative affect. In Study 5, some participants reasoned that following situations which provide readily available counterfactuals individuals might feel more negative about the outcome of the event and these feelings might motivate them to deceive. In Chapter 4 we wanted to further investigate these results and directly assess whether affect could be an underlying mechanism for the relationship between counterfactuals and deception. Results from Study 8 showed that after engaging in downward counterfactual thoughts people deceived more than when considering factual information. However this finding was not related to the way in which counterfactuals made people feel about the events in the scenario. Furthermore, results from Study 9 showed that when counterfactuals did lead to different emotional reactions this difference was not reflected in subsequent deception generation. Therefore, we did not find counterfactually derived affect to influence lying.

An alternative explanation for the link between counterfactuals and deception could relate to the fact that both processes rely on executive functions such as cognitive flexibility and inhibition (Drayton et al., 2011; Gombos, 2006). In study 1 we attempted to assess this potential mechanism by using an individual differences approach and measuring whether the generation of counterfactual thoughts and lies relied on similar executive function skills. We found no evidence for such a hypothesis. We readdressed this hypothesis in Study 11 where we compared counterfactual and deceptive generations between Parkinson's disease patients, a population impaired in executive function and a control sample. We found that impairment in cognitive flexibility did not lead to impairment in either ability. These results contrast with previous findings (Abe et al., 2009; McNamara et al., 2003) and need to be interpreted in relation to the characteristics of the populations assessed as well as the methodology used. Results form Study 11 also revealed an association between counterfactuals and deception, however as opposed to our previous study this relationship was a negative one. We suggest that this is most likely due to the methodology used but these findings do highlight potential situations under certain which the relationship might be moderated and future studies should investigate this thoroughly.

Overall, we did not find either affect or executive functions to be the explanatory mechanism for the relationship between counterfactuals and deception. Previous research has suggested that observing counterfactuals might influence deception because they allow for self-justifications. Namely, Shalvi et al (2011) found that the dishonesty following the observation of desired counterfactuals is judged as less unethical than dishonesty which does not follow the observation of the counterfactual. This finding lead to the suggestion that observing desired

counterfactuals justifies lying. Self-justification might be the psychological mechanism underlying some of our own findings, particularly ones in Study 7. However, in most of our own studies participants had to mentally simulate counterfactuals rather than observe them, and therefore we would like to propose that before enabling a self-justification, counterfactuals motivate deception by increasing the general availability of thoughts as to how things could have happened differently. We consider the availability of alternatives argument and the self-justification argument to be complimentary. In fact, empirical evidence suggests a positive relationship between the number of available alternatives and self-justifications. For example, Schweitzer and Hsee (2002) showed that individuals find it more acceptable to lie when they have a wider range of alternatives available to choose from.

Overall results in this thesis establish a direct relationship between counterfactual thinking and deception. Additionally by investigating the mental simulation of counterfactuals we also highlight that availability of information could be a potential mechanism for this relationship. Next we will emphasize the way in which these findings contribute to the counterfactual thinking and deception literature.

## Implications for counterfactual thinking

The dominant view of counterfactual thinking argues that these thoughts help manage and improve future behaviour by highlighting how things could have been better (Epstude & Roese, 2008; Roese & Epstude, 2017). However, this view has been questioned by evidence suggesting that most spontaneous counterfactuals tend to focus on features outside of our control which do not allow us to improve on the

past (Ferrante et al., 2013; Mercier et al., 2016; but see Roese et al., 2017). This suggests that counterfactuals might serve other functions. Our findings suggest that one potential role of counterfactual thoughts is to allow us to gain personal advantages and avoid costs by enabling us to manipulate others. Given that such behaviour can often lead to substantial social and legal damages one might consider that counterfactuals thus enable maladaptive behaviour, a perspective which highlights the 'dark side' of these thoughts. But lying is not necessarily bad. Indeed the most prevalent lies are white lies which are meant to maintain inter-personal relationships by saving-face and not hurting people's feelings (Serota et al., 2010). Deception is in fact regarded as survival skill for human beings (Brytting, Minogue, & Morino, 2011; Riggio, Tucker, & Throckmorton, 1987). Lies therefore have a social function and accordingly counterfactuals allow us to fulfil such a social function. Overall regardless of the perspective one takes on deception, the fact that counterfactuals enable the generation of lies allows us to understand that these thoughts serve a function in social communication.

## Implications for deception

Deception involves multiple cognitive processes, and the comprehensive identification of these processes has been highlighted as a crucial research aim for the field (Gamer & Ambach, 2014). We wished to address this aim by providing much needed empirical evidence to suggest counterfactual thinking is one of these mechanisms. In doing so, the current research also emphasizes the potential role of individual differences and situational factors in deception.

Firstly, our results show that individual differences in the ability to think counterfactually can impact the tendency and ability to deceive. These results allow

us to further understand the characteristics of a frequent and effective liar, a topic which has been largely overlooked in deception research. Vrij et al. (2010) are the only ones to have listed the characteristics of a good liar. For example they state that a good liar needs to be a rapid thinker, have a good memory and be well prepared. Our research suggests that we can make a novel contribution to this list by adding that good liars need to be skilled at counterfactual thinking. Alternatively, it could also be that counterfactual thinking is an important ingredient for the characteristics already present in the list. For example it might be that in order to be well prepared one needs to be able to think about alternatives to past events they plan to lie about. Such insights are necessary as research suggests that in order to detect deception it might be more useful to ascertain what characterizes a good liar rather than what characterizes a good detector of lies (Bond & DePaulo, 2008). Secondly, we contribute to the literature by further highlighting the way in which characteristics of a situation can allow people to behave unethically. Namely, our results suggest that situations which provide readily available alternatives can compromise our moral compass. This evidence further contributes to the idea that our morality is dynamic and malleable rather than a stable trait (Monin & Jordan, 2009).

Together these findings also have the potential to contribute to the debate concerning the need to build models of deception which take into account both individual and situational characteristics (Gamer & Ambach, 2014). In particular our results give rise to specific hypotheses which have the potential to deliver a more nuanced view of the cognitive load approach. Cognitive load paradigms are based on evidence suggesting that lying is cognitively more demanding than truth-telling and argue that increasing cognitive load even further can help differentiate between truth-tellers and liars (Vrij, Fisher, Mann, & Leal, 2008). But whilst these

approaches appear to be effective at identifying cues to deception, they have been criticised for not identifying individual processes involved in the generation of lies. This is problematic as it does not allow us to understand under which conditions these approaches might be most useful (Blandón-Gitlin, Fenn, Masip, & Yoo, 2014). This is important given that lying is not always more cognitively demanding than telling the truth (Burgoon, 2015). Whilst we did not directly test whether the ability to engage in counterfactual thinking diminishes the cognitive load associated with deception, evidence showing that activation of previous knowledge can make lying cognitively easier suggests that this is a possibility (Falkiewicz et al., 2015). This coupled with the fact that counterfactuals can provide an advantage to the ability to lie should lead to questions about the way in which such thinking can moderate the efficiency of cognitive load approaches. More specifically, we want to suggest that future work should seek to address how efficient such paradigms are in situations in which alternative scenarios are highly available. Furthermore, it could also be worth investigating whether thinking about alternative possibilities can be used by liars as a countermeasure in order to evade detection.

## Methodological considerations

Methodological issues have been considered within each individual chapter, for example, in Chapter 4 we suggested that future studies should seek and understand whether affect might underlie the relationship between counterfactuals and deception by using behavioural tasks which would allow individuals to report on experienced affect. Such limitations must indeed be addressed in order to further our understanding of the relationship between counterfactual thinking and deception. In this section however we will aim to specify a few methodological considerations that

cut across all of our studies and thus emphasize the potential boundaries for the interpretation of our results.

**Measurement of counterfactual thinking**

In this thesis we developed our original methodology which allowed for the direct assessment of the counterfactual thoughts that participants generated. This helped us improve on previous research which only inferred counterfactual thinking through experimental manipulation of die rolls (Shalvi et al., 2011). However, the use of scenario-based measures for the assessment of counterfactual thinking has several drawbacks. Most limitations are related to the fact that participants had to imagine themselves in these situations and did not experience the events themselves. Firstly, this is notable because, although most participants could easily and readily engage with the scenarios, as evidence by their inclusion of names and personal details, we did not directly test to what extent this was the case. Thus we are unable to state whether one's ability to engage with the scenario can moderate our results. Secondly, the imagination of the scenario itself could have added an additional cognitive demand. This could have in theory limited participants' capacity to come up with subsequent lies thus underestimating the relationship between counterfactuals and deception. However this is less likely to be the case given that executive functions were not related to the generation of deceptive responses. Lastly, and most importantly, it has been shown that reflection on impersonal and non-autobiographical events tend to generate different type of counterfactual thoughts in comparison to reflection on past personal events (Girotto et al., 2007; Pighin et al., 2011). Additionally the generation of counterfactuals has been found to be influenced by the degree to which an individual feels personally involved in the

outcome. For example, when participants read a scenario in which a culprit is similar to them, they tend to deflect the blame from the culprit by undoing aspects relating to the situation rather than undoing the culprit's actions (Roese, Black, & Olson, 1994). This might be a limitation given that in most of our studies participants had to generate counterfactuals about events they imagined rather than personally experienced. Whilst participants could have still used their auto-biographical knowledge in order to respond to the scenarios the fact that they did not personally experience the events could have had an impact on the way they responded. This could have an implication for our interpretation that counterfactuals influence deception through a content-neutral pathway. In particular, it could be argued that counterfactuals derived from personal experienced will emphasize different information which be might be more pertinent to the construction of a lie and thus transfer more easily to the content of the lie itself. Under such conditions counterfactuals would influence deception through the content-specific pathway. Whilst this potential hypothesis does not undermine the findings in this thesis, it does provide a direction for future research which could help us understand more about the way in which counterfactuals influence the generations of lies. However, it is also worth mentioning that in Study 7 we did find that participants lied more when counterfactuals were more readily available, suggesting that our results hold in behavioural studies where individuals personally experienced the counterfactual nature of a situation.

**Measurement of deception**

Apart from being able to directly quantify the number of lies participants generated, our methodology also allowed us to assess the spontaneous generation of

lies. This contrasts with most deception paradigms which require participants to lie after being instructed to do so thus not allowing for personal choice (Wright et al., 2012). However, it is important to consider whether our tasks matched the demands of lying outside of the laboratory.

Firstly, one might argue our participants were not strongly motivated to lie. The scenarios we used were chosen so that we can include both high and low-stakes given that such circumstances motivate deception differently (ten Brinke & Porter, 2012). However, given that these scenarios were hypothetical there was no real punishment for either the choice to either be truthful or to be deceptive. As such, the scenario-based methodology potentially diminished any risk associated with lying as well as any potential for feelings of guilt and anxiety which would normally accompany real life deception (Caso, Gnisci, Vrij, & Mann, 2005). The hypothetical nature of our scenarios could indeed be why we did not observe an effect of counterfactually derived affect in our studies from Chapter 4.

The methodology we used in Study 6 and 7 addressed the motivation issue by providing personal rewards for the decision to lie. However lies were sanctioned as the participants were explicitly told they could lie and thus the decision to lie would not have transgressed moral rules in the same way lying in the real world usually would. Having said that, in Study 7 participants knew that their decision to lie would have affected someone else and thus this would not have entirely removed moral transgression but rather diminished it. This coupled with the fact that others have supported the use of sanctioned lies to examine interpersonal deception means that this might not been a limitation in our research (Feeley, 1996).

Lastly, it is also worth considering that although responses to our scenarios required the imagination of a conversational exchange, most of our studies did not involve actual social interaction (apart from Study 6). This contrasts with real-life deception which often involves a dynamic interaction (Buller & Burgoon, 1996) but is typical of most paradigms assessing deception (Bond & DePaulo, 2006). The absence of social interaction might have diminished the cognitive load associated with lie generation as well as participants' need to adapt their dishonest statements in response to further questioning (Burgoon, 2015). These are also potential reasons for why we did not find a clear association between lying and executive functions in Study 1 or Study 11.

These considerations confer an understanding of the limits of our experimental approach. Nonetheless we believe that this approach was appropriate given our overall aim as it allowed us to test the role of counterfactual thinking in deception in a controlled way free from the noise that real life situations might bring. Independent of limitations related to social interaction, moral transgression and motivation, in order to lie, our participants still needed to come up with statements which offered an alternative to the truth as presented in the scenarios. This is crucial given that this process is potentially necessary for most lies (Debey et al., 2014) and that our aim was to specifically investigate the relationship between this ability and the ability to come up with alternatives to the past. However, the methodological limitations highlighted here should be taken into account in the future, especially if trying to understand the relationship between counterfactual thinking and the quality of a lie.

## Future directions

As highlighted above, the primary focus of the current thesis was to establish a link between counterfactual thinking and deception at a general level. In doing so we attempted to average across different types of situations in which deception is likely to occur. Future research should seek to understand the relationship by evaluating the different characteristics of counterfactuals and lies in relation to one another.

One element that could be considered is the length of the deceptive response. A benefit of the methodology we used in the current thesis is that most of the questions we asked in order to assess deception were open ended and thus did not elicit simple binary responses (apart from Study 7). However, whilst some participants provided simple answers such as ‘I don’t know anything about this’ others generated more complex lies which explained an alternative version of events in detail. The construction of complex lies calls on additional cognitive processes in comparison to simpler deceptive statements (DePaulo et al., 2003). As such, it would be interesting to assess whether counterfactual thinking is more useful for the construction of complex lies.

The structure of counterfactual thoughts could also be considered in future research. Thoughts about alternatives to the past can be constructed either by subtracting elements from what actually happened (If only it hadn't rained today, I would not have gotten wet”) or by adding them (e.g. “If only I had an umbrella, I would not have gotten wet) (Roese & Olson, 1993) . Such thoughts have been found have differential effects on mood and behavioural regulation (Markman et al., 2007; Roese, 1994) and thus it might be valuable to assess whether they can also have

differential effects on lie construction. This examination would be particularly relevant if performed in relation to different types of lies such as lies which are designed to mislead by omitting relevant details (i.e. omissions) or lies which convey impressions which exceed the truth (i.e. exaggerations) (DePaulo et al., 1996).

In the current thesis we approached the relationship between counterfactuals and deception by considering the impact of counterfactual thinking on subsequent deception, however it is also plausible to assume that the need to deceive drives the consideration of alternative possibilities. Such a hypothesis would be in line with evidence which suggests that the construction of counterfactual thoughts is dependent on situational need (Roese & Epstude, 2017). Thus, if deception is needed in order to achieve a desired goal, one might spontaneously activate counterfactuals with the aim of achieving such a goal. One way to investigate this would be to manipulate the motivation to lie and subsequently assess its effect on the generation of counterfactual thoughts. If the relationship is bi-directional we would expect people to generate more counterfactuals in situations where the motivation to lie is increased. Such evidence would still be in support our interpretation that counterfactual thinking is an important process in the generation of lies.

## Concluding remarks

This thesis commenced with a quote which emphasized that imagination can mislead us. The results in this thesis add a different dimension to the statement by showing that counterfactuals can also be used to mislead others. The studies in this thesis constitute the first in-depth empirical research to directly assess counterfactuals and deception in relation to one another. Overall the findings allow us to understand more about the ways in which we use counterfactual thoughts and

also the way in which we come to generate lies. We believe this work helps shed more light onto processes we use on a day-to-day basis and also provides the impetus for a new stream of research.

## Appendix 2A

## Scenarios used in Study 1

### Moving town scenario

You're moving house to start a new job in a different city. The night before you leave, you write down your feelings about the move in your diary:

"I've got mixed feelings about moving to a place where I hardly know anyone. I'm sure it will be easy to settle into the new town – I've never had any trouble making new friends and I have always been very comfortable around strangers."

A lot happens in your first two weeks in the new town. During your first week at work, a staff dinner is held. You decide to go because you want to get to know your colleagues. You enjoy the evening and meet a lot of people. That weekend, your next-door neighbours invite you to a party. Most of the people who live on your road will be there. You promise your neighbour you will attend but then at the last minute you decide you would rather go to the cinema on your own as you really want to the new movie being shown. The next week you decide to ring an old friend who lives in the town and ask him to show you around. You arrange to go out with him the following evening and he introduces you to a lot of his friends. A few days later, a colleague tells you there's a membership vacancy at her sports club. You think joining would be a good way to meet people, but then you decide to spend the money on a new stereo instead. Six weeks after the move, things have turned out nothing like you had expected. You haven't made any real friends in the new town and you feel very lonely and isolated. You are very upset and very surprised.

### Car crash scenario

It is Friday evening, you have had a really busy week and are tired, but you have little food left so you decide to go do some shopping.

You chose to go to the supermarket outside the town rather than your corner shop as you want to buy quite a few things. When you get to the supermarket you realise that everybody else must have had the same plan as you as the shop is very busy. You get everything you need but this takes a long time and when you finish you are even more worn out. Driving back the usual way would take too much time so you decide to drive home down a new route. The route you chose has heavy traffic so you end up taking even longer to get home. The roads on this route are also narrower so you really have to concentrate whilst driving. You become very tired and you feel your attention is slipping. When you finally get closer to home, your phone rings. As it is one of your relatives you decide to answer and put the phone on loud speaker. As you try to pay attention to what they are saying you end up driving too close to the cars that are parked on the road. You scratch someone's car and clip their wing mirror.

## Appendix 2B

## Questions asked to assess spontaneous and cued generations of counterfactuals and deception in Study 1

**Spontaneous counterfactuals:**

Considering the previous scenario please imagine the thoughts and feelings you would have and write them down as a diary entry.

**Cued counterfactuals:**

When you think back on past events, you may think about how things might have happened differently, for example, “if only…” or “what if…”. I would now like you to think about the scenario that you have just read. In the space below, please write any thoughts about how things might have happened differently. You may look back at the scenario as often as you wish.

**Spontaneous deception**

**Moving town scenario indirect deception question:**

Imagine that you bump into your neighbour for the first time after the party. Please write down everything that you think you would say to him/her.

**Moving town scenario direct deception question:**

The neighbour then specifically asks you: ‘You promised to attend the party but you never showed up, what happened? How do you reply?

**Car incident scenario indirect deception question:**

Imagine that after the accident, you immediately look around to see whether anyone has seen you. The street is quiet and you are sure nobody has seen you. As you are tired and very close to home you just decide to drive past. The next day the police are at your door, they have no idea that you are the culprit, but a neighbour has put in a complaint and they are trying to investigate what happened. Please write down everything that you think you would say to the police.

**Car incident scenario direct deception question:**

They then specifically ask you: what do you know about the damaged car? What do you say?

## Cued deception

### Moving town scenario cued deception question:

Imagine you didn't want your neighbour to know you went to the cinema instead of attending the party, what would you say? Please list some different possible things that you could tell him / her.

### Car incident scenario cued deception question:

Imagine you didn’t want the police to know that you are the culprit. Please list some different possible things that you could tell them.

## Appendix 2C

## Coding scheme used for coding counterfactual thinking

### *Include*

Imagining something that didn't happen, e.g.,

- If only/If I had gone to on the school trip
- I should have gone on the school trip
- I could have …
- I wish I had gone on the school trip
- I regret …
- I should have done A and B counts as 2
- If x had happened then A and B would have happened counts as 2 when A and B are different
- If x or y had happened then A and B would have happened counts as 3 when A and B are different
- Present situations, e.g., If I were living elsewhere…
- In general, things should be this way ….

They can also include more general events

- I should have done things differently
- I wish I had never moved

Explanations for things that did happen

- why didn't I go on the school trip
- I was stupid not going on the trip

Just mentioning the outcome

- I thought things would have been better

### *Don't include*

Things are better this way

Repetitions and variations of things that they have already mentioned

- I wish I had gone on the school trip and not bought the stereo (only counts as 1 not 2)
- Repetition of something already mentioned

## Appendix 2D
## Coding scheme used to code deceptive responses

Definition of lies:

*'Most people think a lie occurs any time you intentionally try to mislead someone. Some lies are big while others are small; some are completely false statements and others are truths with a few essential details made up or left out. Some lies are obvious, and some are very subtle. Some lies are told for a good reason. Some lies are selfish, other lies protect others. We are interested in all these different types of lies.'*

- Each lie no matter how big or small will be given a score of 1.
- If participant repeats the same lie twice only score this a 1
- Responses related to feelings and emotions (e.g. social anxiety) are classed as lies as long as the subject doesn't state (this is true or honestly in the paragraph) and as long as the paragraph doesn't include a confession (i.e. I have been to the movies; or It was me who hit the car).
- Each paragraph and sentence may contain more than one lie.
- In a paragraph multiple lies might be separated by the symbol '/' but they might also be included in the same sentence. for example the sentence I know nothing about the accident, good luck with your investigation, includes <u>two lies</u>: one is the denial of knowledge and the other one relates to 'wishing good luck'.
- Questions are classed as lies if the subject already knows the answer but is trying to mislead by asking the question. For example: Where was the car? Or 'Did it happen on this road'
- A paragraph can include a lie even if it includes a confession. E.g. It was me who hit the car but I left a note in the windshield.
- If you are not sure how to code a lie put a ? in the box and discuss later.

## Appendix 2E
## Scenarios used in Studies 4 and 5

**1.** Adele and Rachel are taking part in an experiment. They are given a die and told that if they roll a six three times they will win ten pounds. The experimenter cannot see the actual die rolls. Adele rolls two sixes, Rachel rolls one six.

Counterfactual: Who spends more time thinking about how things could have been different?

Deception: Who is more likely to lie to the experimenter about the number of sixes they threw?

**2.** Alison and Sarah's parents ask them to be home in time for an important delivery after work. On her way, Alison stops off for a coffee for half an hour, something she usually does on her way home. Sarah also stops off for a coffee for half an hour on her way home although this is something she normally never does. Both of them arrive home just before the delivery was due only to discover that it arrived 5 minutes early and they missed it.

Counterfactual: Who spends more time thinking about how things could have been different?

Deception: Who is more likely to lie to their parents about stopping on the way home?

**3.** Jenny and Jacqui want to impress their friends by winning a game. In order to win the game they need to achieve a total score of 20 points. Jenny loses the game by 1 point. Jacqui loses the game by 10 points. Jenny and Jacqui's friends don't know what the score actually was.

Counterfactual: Who spends more time thinking about how things could have been different?

Deception: Who is more likely to lie to their friends about winning the game?

**4**. The day before their friend's birthday Jack and Ed remember that they forgot to send them a birthday card. The last official collection of the post for the day is at 5pm. Both run to the post office but they get there late at 5:15. Jack finds out that the mail had been collected in time 15 min earlier. Ed finds out that the post was collected late, just one minute before they arrived.

Counterfactual: Who spends more time thinking about how things could have been different?

Deception: Who is more likely to lie to their friend about posting the letter on time?

**5**. John crashes his dad's car while driving on his usual way home. Bob crashes his dad's car whilst trying a new way home.

Counterfactual:  Who spends more time thinking about how things could have been different?

Deception: Who is more likely to lie to their dad about crashing the car?

**6**. Steve and Aaron are asked by a friend to safeguard a valuable object for them. On his way home Steve gets off the bus at his usual bus stop but once the bus is gone realises that he forgot the object on the bus. On his way home Aaron gets off at a different stop from the one he usually gets off at and once the bus is gone realises that he forgot the valuable object on the bus.

Counterfactual:  Who spends more time thinking about how things could have been different?

Deception: Who is more likely to lie to their friend about losing the item?

## Appendix 3A
## Episodic memory items used in Study 6

1. Describe what you did last weekend?
2. Describe what you did for your last birthday
3. Describe what you did during your last summer vacation
4. Describe what you did last night
5. Describe your home town
6. Describe what you did for Christmas last year
7. Describe your favourite restaurant
8. Describe what you did yesterday morning
9. Describe your favourite movie
10. Describe what you did for New Year's Eve

## Appendix 3B
## Response sheets used for Sender and Receiver roles in Study 6

**Sender Role:**

After providing a response to the topic given on screen please circle T for truth and L for Lie depending on the way in which you answered the statement

| | **1** | | **2** | | **3** | | **4** | | **5** | | **6** | | **7** | | **8** | | **9** | | **10** | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | T | L | T | L | T | L | T | L | T | L | T | L | T | L | T | L | T | L | T | L |

**Receiver Role:**

After the participant has provided a statement please note down whether you think they have told the truth by writing a T or whether you think they told a lie by writing an L for the corresponding trial number.

| | **Trial Number** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Participant number** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 1 | | | | | | | | | | |
| 2 | | | | | | | | | | |
| 3 | | | | | | | | | | |
| 4 | | | | | | | | | | |

## Appendix 3C
## Instructions provided to participants in Study 7

Please read carefully**, you may earn an extra ½ point** depending on the decisions you and the other participants make!

During this study you will play a game with two others in this room. You will never know who the others are, and they will never know who you are. If you win the extra ½ will be added to your account after the end of the study.

Due to the seat you have chosen, you will be **person A.** At the beginning of the game you will randomly be matched with two other people called person B and person C. The aim of the game is for you to gain more points than person B. Only ONE of you will get to win the extra ½ point. The winner will be the person who won the most points. Points are earned in the following way:

In each of the 24 rounds you will be assigned a number. In order to gain a point, you need to spin a slot machine and get the two numbers on the machine to add up to the assigned number. For example, if you are assigned a 2, you need to spin either a 0 and a 2 OR a 1 and a 1. If the numbers that you spin don't add up to the assigned number, you get no points. Also, if the slot machine stops in between numbers, only the smaller number of the two will count. In each round the computer will display the sum of the numbers you scored, and colours will indicate whether the round was lost or won, red for a loss, green for a win.

At the end of each round, you will be asked to report whether you won or lost the round. CRUCIALLY, **clicking 'Y' (for yes), will ALWAYS gain you a point** (even if the two numbers do not add up to the assigned number). However, by clicking 'Y' when the numbers do not add up to the assigned number you will be **negatively impacting person C**.

This is because person C will get to win ½ point only if they correctly identify the truthfulness of your messages. If you report a win and this is false, and person C chooses to class it as truthful, they lose a point. From our previous studies we know that the only sure way to help person C win is to be truthful, this is because in most cases person C will most likely start with the assumption that the other players will tell the truth. However, this could affect your own win.

In summary, you have to compete against player B, and win more points in order to win the extra participation point, however you can also help or hinder participant's C chances of winning depending on the messages you choose to send.

## Appendix 4A
## Scenarios used in Study 8

### Moving town scenario

You recently moved house to start a new job in a different city where you don't have any friends. A few weeks after your move, your next-door neighbour invites you to a party. Most of the people who live on your road will be there. You promise your neighbour you will attend but then at the last minute you decide you would rather go to the cinema on your own as you really want to see the new movie being shown. On your return home you feel lonely and isolated

### Car incident scenario

Friday, after a busy week at work you decide to go shopping. On your way back from the supermarket you become very tired and you feel your attention is slipping. When you finally get closer to home, your phone rings. As it is one of your relatives you decide to answer. As you try to pay attention to what they are saying you end up driving too close to the cars that are parked on the road. You scratch someone's car and clip their wing mirror.

### Delivery scenario

Your housemate is going on an overnight trip and will not be home this evening. They have an important package that is being delivered later in the day and they ask you to be home in time to collect it. You agree. You leave work a bit early and realise you have time to stop for a drink at a coffee shop on the way back. You arrive home later than you had planned and realised you have missed the delivery.

## Appendix 4B
## Supplementary results relating to rating of affect in Study 10

When analysing each moral affect item separately, all ratings were heightened for the 'car crash' scenario in comparison to the 'moving town' scenario. Descriptive statistics are provided in Table 4B.1. There was no main effect of condition and only the analysis concerning regret revealed an interaction between scenario and condition $F$ (2, 147) = 3.43, $p$ =.035, $\eta_p^2$ =.05.

Simple main effects analyses for the interaction regarding **regret** revealed that there was a difference between scenarios with participants felt more regretful in response to the car scenario than in response to the moving town scenario. Descriptive statistics for regret ratings are depicted graphically in Figure 4B.1.

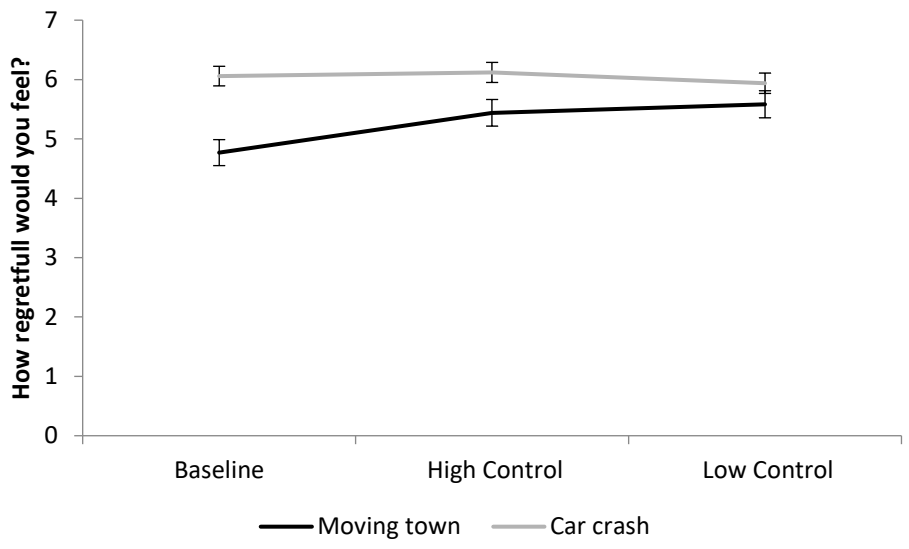


**Figure 4B.1** Mean rating of regret for the two scenarios and three experimental conditions, (higher score indicate increased regret).

This difference was however prevalent in the baseline condition ($p <. 001$) and the high control condition ($p = .009$) but not the low control condition ($p = .179$). Mean comparison between conditions as a function of scenario revealed that there were no statistical differences between conditions for the car crash scenario. However when looking at the moving town scenario, there were statistically significant differences in terms of regret between participants in the baseline and high control condition ($p =.034$) and baseline and low control condition ($p =.011$), but no differences between high and low control conditions ($p = .654$).

Table 4B.1

*Means and standard deviations for all moral affect items for each scenario across conditions (Study 10)*

| Scenario | Emotion | Condition | | |
|---|---|---|---|---|
| | | Baseline *M (SD)* | High Control *M (SD)* | Low Control *M (SD)* |
| Moving town scenario | Guilt | 4.06 (1.61) | 4.44 (1.43) | 4.48 (1.86) |
| | Shame | 3.39 (1.90) | 3.96 (1.69) | 3.81 (1.76) |
| | Regret | 4.77 (1.86) | 5.44 (1.47) | 5.58 (1.33) |
| | Disappointment | 4.31 (1.91) | 4.66 (1.73) | 4.79 (1.75) |
| Car crash scenario | Guilt | 6.04 (1.27) | 6.22 (1.17) | 6.35 (1.21) |
| | Shame | 5.62 (1.32) | 5.68 (1.50) | 5.73 (1.48) |
| | Regret | 6.06 (1.21) | 6.12 (1.26) | 5.94 (1.10) |
| | Disappointment | 5.73 (1.25) | 5.74 (1.59) | 5.29 (1.61) |

When evaluating ratings of **guilt**, analyses showed an effect for scenario $F$ (1, 147) = 206.29, $p$ <.001, $\eta_p^2$ =.58, with participants feeling more guilt in the 'car crash' scenario ($M$ = 6.20, $SD$ =1.22) than the 'moving town scenario' ($M$ = 4.32, $SD$ =1.64). There was no main effect of condition $F$ (2, 147) = 1.30, $p$ =.275, $\eta_p^2$ =.02, and no significant interaction between scenario and condition $F$ (2, 147) = .20, $p$ =.818, $\eta_p^2$ <.01.

In terms of **shame**, analyses showed a main effect of scenario $F$ (1, 147) = 156.91, $p$ <.001, $\eta_p^2$ =.52, with participants feeling more shame in response to the 'car crash' scenario ($M$ = 5.67, $SD$ =1.43) than the 'moving town scenario' ($M$ = 3.71, $SD$ =1.79). There was no main effect of condition $F$ (2, 147) = .89, $p$ =.415, $\eta_p^2$ =.01, and no significant interaction between scenario and condition $F$ (2, 147) = .93, $p$ =.397, $\eta_p^2$ <.01.

When analysing **disappointment**, we found a main effect of scenario $F$ (1, 147) = 34.90, $p$ <.001, $\eta_p^2$ =.19, with participants feeling more disappointed in response to the 'car crash' scenario ($M$ = 5.59, $SD$ =1.49) than the 'moving town scenario' ($M$ = 4.58, $SD$ =1.75). There was no main effect of condition $F$ (2, 147) = .31, $p$ =.733, $\eta_p^2$ < .01, and no significant interaction between scenario and condition $F$ (2, 147) = 2.51, $p$ = .085, $\eta_p^2$ <.03.

## References


Abe, N. (2009). The neurobiology of deception: evidence from neuroimaging and loss-of-function studies. *Current opinion in neurology, 22*(6), 594-600. doi: 10.1097/WCO.0b013e328332c3cf

Abe, N. (2011). How the Brain Shapes Deception: An Integrated Review of the Literature. *Neuroscientist, 17*(5), 560-574. doi: 10.1177/1073858410393359

Abe, N., Fujii, T., Hirayama, K., Takeda, A., Hosokai, Y., Ishioka, T., . . . Mori, E. (2009). Do parkinsonian patients have trouble telling lies? The neurobiological basis of deceptive behaviour. *Brain, 132*(5), 1386-1395. doi: 10.1093/brain/awp052

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. doi: 10.1093/cercor/bhn037

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. doi: 10.1093/cercor/bhi097

Allen, M. S., Greenlees, I., & Jones, M. V. (2014). Personality, counterfactual thinking, and negative emotional reactivity. *Psychology of Sport and Exercise, 15*(2), 147-154. doi: 10.1016/j.psychsport.2013.10.011

Alloway, T. P., McCallum, F., Alloway, R. G., & Hoicka, E. (2015). Liar, liar, working memory on fire: Investigating the role of working memory in childhood verbal deception. *Journal of Experimental Child Psychology, 137*, 30-38. doi: 10.1016/j.jecp.2015.03.013

Bacon, A. M., Walsh, C. R., & Briazu, R. A. (under review). Personality and thinking about what might have been.

Barbey, A. K., Krueger, F., & Grafman, J. (2009). Structured event complexes in the medial prefrontal cortex support counterfactual representations for future planning. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 364*(1521), 1291-1300. doi: 10.1098/rstb.2008.0315

Barnacz, A. L., Johnson, A., Constantino, P., & Keenan, J. P. (2004). Schizotypal personality traits and deception: The role of self-awareness. *Schizophrenia Research, 70*(1), 115-116. doi: 0.1016/j.schres.2003.12.006

Bassarak, C., Leib, M., Mischkowski, D., Strang, S., Glöckner, A., & Shalvi, S. (2017). What Provides Justification for Cheating—Producing or Observing Counterfactuals? *Journal of Behavioral Decision Making*. doi: 10.1002/bdm.2013

Baumeister, R. F., Stillwell, A. M., & Heatherton, T. F. (1995). Personal narratives about guilt: Role in action control and interpersonal relationships. *Basic and*

*Applied Social Psychology, 17*(1-2), 173-198. doi: 10.1080/01973533.1995.9646138

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. doi: 10.1016/j.tsc.2012.10.003

Beck, S. R., & Guthrie, C. (2011). Almost thinking counterfactually: Children's understanding of close counterfactuals. *Child Development, 82*(4), 1189-1198. doi: 10.1111/j.1467-8624.2011.01590.x

Beck, S. R., Riggs, K., & Burns, P. (2011). Multiple Development in Counterfactual Thinking. In C. T. Hoerl, T. McCormack & S. R. Beck (Eds.), *Understanding Counterfactuals, Understanding Causation: Issues in Philosophy and Psychology*: Oxford University Press.

Beck, S. R., Riggs, K. J., & Gorniak, S. L. (2009). Relating developments in children's counterfactual thinking and executive functions. *Thinking & Reasoning, 15*(4), 337-354. doi: 10.1080/13546780903135904

Beck, S. R., Riggs, K. J., & Gorniak, S. L. (2010). The effect of causal chain length on counterfactual conditional reasoning. *British Journal of Developmental Psychology, 28*(3), 505-521. doi: 10.1348/026151009X450836

Beck, S. R., Robinson, E. J., Carroll, D. J., & Apperly, I. A. (2006). Children's thinking about counterfactuals and future hypotheticals as possibilities. *Child Development, 77*(2), 413-426. doi: 10.1111/j.1467-8624.2006.00879.x

Beldarrain, M. G., Garcia-Monco, J. C., Astigarraga, E., Gonzalez, A., & Grafman, J. (2005). Only spontaneous counterfactual thinking is impaired in patients with prefrontal cortex lesions. *Cognitive Brain Research, 24*(3), 723-726. doi: 10.1016/j.cogbrainres.2005.03.013

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development, 81*(6), 1641-1660. doi: 10.1111/j.1467-8624.2010.01499.x

Blanchard-Fields, F., Mienaltowski, A., & Seay, R. B. (2007). Age differences in everyday problem-solving effectiveness: Older adults select more effective strategies for interpersonal problems. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 62*(1), 61-P64. doi: 10.1093/geronb/62.1.P61

Blandón-Gitlin, I., Fenn, E., Masip, J., & Yoo, A. H. (2014). Cognitive-load approaches to detect deception: searching for cognitive mechanisms. *Trends in Cognitive Sciences, 18*(9), 441-444. doi: 10.1016/j.tics.2014.05.004

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, 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

Brytting, T., Minogue, R., & Morino, V. (2011). *The anatomy of fraud and corruption: Organizational causes and remedies*: Gower Publishing, Ltd.

Buck, M. L., & Miller, D. T. (1994). Reactions to incongruous negative life events. *Social Justice Research, 7*(1), 29-46. doi: 10.1007/BF02333821

Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory, 6*(3), 203-242. doi: 10.1111/j.1468-2885.1996.tb00127.x

Burgoon, J. K. (2015). When is Deceptive Message Production More Effortful than Truth-Telling? A Baker's Dozen of Moderators. *Frontiers in Psychology, 6*. doi: 10.3389/fpsyg.2015.01965

Burgoon, J. K., Stoner, G., Bonito, J. A., & Dunbar, N. E. (2003). *Trust and deception in mediated communication.* Paper presented at the System Sciences, 2003. Proceedings of the 36th Annual Hawaii International Conference on System Sciences.

Burns, P., Riggs, K. J., & Beck, S. R. (2012). Executive control and the experience of regret. *Journal of Experimental Child Psychology, 111*(3), 501-515. doi: 10.1016/j.jecp.2011.10.003

Byrne, R. M. (2002). Mental models and counterfactual thoughts about what might have been. *Trends in Cognitive Science, 6*(10), 426-431. doi: 10.1016/S1364-6613(02)01974-5

Byrne, R. M. (2005). *The Rational Imagination: How people create alternatives to reality*: MIT Press.

Byrne, R. M. (2016). Counterfactual thought. *Annual Review of Psychology, 67*, 135-157. doi: 10.1146/annurev-psych-122414-033249

Camille, N., Coricelli, G., Sallet, J., Pradat-Diehl, P., Duhamel, J. R., & Sirigu, A. (2004). The involvement of the orbitofrontal cortex in the experience of regret. *Science, 304*(5674), 1167-1170. doi: 10.1126/science.1094550

Caso, L., Gnisci, A., Vrij, A., & Mann, S. (2005). Processes underlying deception: an empirical analysis of truth and lies when manipulating the stakes. *Journal of Investigative Psychology and Offender Profiling, 2*(3), 195-202.

Catellani, P., & Bertolotti, M. (2014). The effects of counterfactual defences on social judgements. *European Journal of Social Psychology, 44*(1), 82-92. doi: 10.1002/ejsp.1995

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. doi: 10.1093/cercor/bhn189

Contreras, F., Albacete, A., Castellví, P., Caño, A., Benejam, B., & Menchón, J. M. (2016). Counterfactual reasoning deficits in schizophrenia patients. *PloS One, 11*(2). doi: 10.1371/journal.pone.0148440

Coricelli, G., Critchley, H. D., Joffily, M., O'Doherty, J. P., Sirigu, A., & Dolan, R. J. (2005). Regret and its avoidance: a neuroimaging study of choice behavior. *Nature Neuroscience, 8*(9), 1255-1262. doi: 10.1038/nn1514

Credé, M., Harms, P., Niehorster, S., & Gaye-Valentine, A. (2012). An evaluation of the consequences of using short measures of the Big Five personality traits. *Journal of Personality and Social Psychology, 102*(4), 874. doi: 10.1037/a0027403

Dalbert, C. (2009). Belief in a just world. In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (pp. 288-297). New York: Guilford Publications.

Davis, C. G., Lehman, D. R., Wortman, C. B., Silver, R. C., & Thompson, S. C. (1995). The Undoing of Traumatic Life Events. *Personality and Social Psychology Bulletin, 21*(2), 109-124. doi: Doi 10.1177/0146167295212002

De Brigard, F., Addis, D. R., Ford, J. H., Schacter, D. L., & Giovanello, K. S. (2013). Remembering what could have happened: Neural correlates of episodic counterfactual thinking. *Neuropsychologia, 51*(12), 2401-2414. doi: 10.1016/j.neuropsychologia.2013.01.015

De Brigard, F., Spreng, R. N., Mitchell, J. P., & Schacter, D. L. (2015). Neural activity associated with self, other, and object-based counterfactual thinking. *Neuroimage, 109*, 12-26. doi: 10.1016/j.neuroimage.2014.12.075

Debey, E., De Houwer, J., & Verschuere, B. (2014). Lying relies on the truth. *Cognition, 132*(3), 324-334. doi: 10.1016/j.cognition.2014.04.009

Debey, E., De Schryver, M., Logan, G. D., Suchotzki, K., & Verschuere, B. (2015). From junior to senior Pinocchio: A cross-sectional lifespan investigation of deception. *Acta Psychologica, 160*, 58-68.

Debey, E., Liefooghe, B., De Houwer, J., & Verschuere, B. (2015). Lie, truth, lie: the role of task switching in a deception context. *Psychological Research, 79*(3), 478-488. doi: 10.1007/s00426-014-0582-4

Debey, E., Ridderinkhof, R. K., De Houwer, J., De Schryver, M., & Verschuere, B. (2015). Suppressing the truth as a mechanism of deception: Delta plots reveal the role of response inhibition in lying. *Consciousness and Cognition, 37*, 148-159. doi: 10.1016/j.concog.2015.09.005

Debey, E., Verschuere, B., & Crombez, G. (2012). Lying and executive control: an experimental investigation using ego depletion and goal neglect. *Acta Psychologica, 140*(2), 133-141. doi: 10.1016/j.actpsy.2012.03.004

Denney, D. R., & Lynch, S. G. (2009). The impact of multiple sclerosis on patients' performance on the Stroop Test: processing speed versus interference. *Journal of the International Neuropsychological Society, 15*(03), 451-458. doi: 10.1017/S1355617709090730

Depalma, M. T., Madey, S. F., & Bornschein, S. (1995). Individual-Differences and Cheating Behavior - Guilt and Cheating in Competitive Situations. *Personality and Individual Differences, 18*(6), 761-769. doi: 10.1016/0191-8869(94)00208-A

DePaulo, B. M., Ansfield, M. E., Kirkendol, S. E., & Boden, J. M. (2004). Serious lies. *Basic and Applied Social Psychology, 26*(2-3), 147-167. doi: 10.1080/01973533.2004.9646402

DePaulo, B. M., & Kirkendol, S. E. (1989). The motivational impairment effect in the communication of deception. In J. C. Yuille (Ed.), *Credibility Assessment* (pp. 51-70): Springer.

DePaulo, B. M., Kirkendol, S. E., Kashy, D. A., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology, 70*(5), 979-995. doi: 10.1037/0022-3514.70.5.979

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129*(1), 74-118. doi: 10.1037//0033-2909.129.1.74

Ding, X. P., Gao, X., Fu, G., & Lee, K. (2013). Neural correlates of spontaneous deception: A functional near-infrared spectroscopy (fNIRS)study. *Neuropsychologia, 51*(4), 704-712. doi: 10.1016/j.neuropsychologia.2012.12.018

Ding, X. P., Wellman, H. M., Wang, Y., Fu, G., & Lee, K. (2015). Theory-of-mind training causes honest young children to lie. *Psychological Science, 26*(11), 1812-1821. doi: 10.1177/0956797615604628

Dirnberger, G., & Jahanshahi, M. (2013). Executive dysfunction in Parkinson's disease: a review. *Journal of Neuropsychology, 7*(2), 193-224. doi: 0.1111/jnp.12028

Donagan, A. (1979). *The theory of morality*: University of Chicago Press.

Drayton, S., Turley-Ames, K. J., & Guajardo, N. R. (2011). Counterfactual thinking and false belief: The role of executive function. *Journal of Experimental Child Psychology, 108*(3), 532-548.

Eddy, C. M., Beck, S. R., Mitchell, I. J., Praamstra, P., & Pall, H. S. (2013). Theory of mind deficits in Parkinson's disease: a product of executive dysfunction? *Neuropsychology, 27*(1), 37. doi: 10.1037/a0031302

Effron, D. A., Miller, D. T., & Monin, B. (2012). Inventing racist roads not taken: The licensing effect of immoral counterfactual behaviors. *Journal of Personality and Social Psychology, 103*(6), 916. doi: 10.1037/a0030008

Ekman, P. (1988). Lying and nonverbal behavior: Theoretical issues and new findings. *Journal of Nonverbal Behavior, 12*(3), 163-175. doi: 10.1007/BF00987486

Epstude, K., & Roese, N. J. (2008). The functional theory of counterfactual thinking. *Personality and Social Psychology Review, 12*(2), 168-192. doi: 10.1177/1088868308316091

Evans, A. D., & Lee, K. (2011). Verbal deception from late childhood to middle adolescence and its relation to executive functioning skills. *Developmental Psychology, 47*(4), 1108. doi: 10.1037/a0023425

Evans, A. D., & Lee, K. (2013). Emergence of lying in very young children. *Developmental Psychology, 49*(10), 1958. doi: 10.1037/a0031409

Evans, A. D., Xu, F., & Lee, K. (2011). When all signs point to you: lies told in the face of evidence. *Developmental Psychology, 47*(1), 39. doi: 10.1037/a0020787

Falkiewicz, M., Sarzyńska, J., Babula, J., Szatkowska, I., Grabowska, A., & Nęcka, E. (2015). Explicit instructions increase cognitive costs of deception in predictable social context. *Frontiers in Psychology, 6*. doi: 10.3389/fpsyg.2015.01863

Faust-Socher, A., Kenett, Y., N, Cohen, O., S, Hassin-Baer, S., & Inzelberg, R. (2014). Enhanced creative thinking under dopaminergic therapy in Parkinson disease. *Annals of Neurology, 75*(6), 935-942. doi: 10.1002/ana.24181

Feeley, T. H. (1996). Exploring sanctioned and unsanctioned lies in interpersonal deception. *Communication Research Reports, 13*(2), 164-173. doi: 10.1080/08824099609362083

Feldman, R. S., Forrest, J. A., & Happ, B. R. (2002). Self-presentation and verbal deception: Do self-presenters lie more? *Basic and Applied Social Psychology, 24*(2), 163-170. doi: 10.1207/153248302753674848

Ferrante, D., Girotto, V., Straga, M., & Walsh, C. (2013). Improving the past and the future: a temporal asymmetry in hypothetical thinking. *Journal of Experimental Psychology: General, 142*(1), 23-27. doi: 10.1037/a0027947

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*(3), 189-198. doi: 10.1016/0022-3956(75)90026-6

Gächter, S., & Schulz, J. F. (2016). Intrinsic honesty and the prevalence of rule violations across societies. *Nature*(531), 496-499. doi: 10.1038/nature17160

Gamer, M., & Ambach, W. (2014). Deception research today. *Frontiers in Psychology, 5*. doi: 10.3389/Fpsyg.2014.00256

Ganis, G., Kosslyn, S. M., Stose, S., Thompson, W. L., & Yurgelun-Todd, D. A. (2003). Neural correlates of different types of deception: An fMRI investigation. *Cerebral Cortex, 13*(8), 830-836. doi: 10.1093/cercor/13.8.830

Gaspar, J. P., Seabright, M. A., Reynolds, S. J., & Yam, K. C. (2015). Counterfactual and Factual Reflection: The Influence of Past Misdeeds on Future Immoral Behavior. *The Journal of Social Psychology, 155*(4), 1-11. doi: 10.1080/00224545.2015.1015477

Gaut, B. (2003). Creativity and imagination. In B. Gaut & P. Livingston (Eds.), *The Creation of Art: New Essays in Philosophical Aesthetics* (pp. 148-173): Cambridge University Press.

Gavanski, I., & Wells, G. L. (1989). Counterfactual processing of normal and exceptional events. *Journal of Experimental Social Psychology, 25*(4), 314-325. doi: 10.1016/0022-1031(89)90025-5

Giammarco, E. A., Atkinson, B., Baughman, H. M., Veselka, L., & Vernon, P. A. (2013). The relation between antisocial personality and the perceived ability to deceive. *Personality and Individual Differences, 54*(2), 246-250. doi: 10.1016/j.paid.2012.09.004

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. doi: 10.1037/a0026406

Girotto, V., Ferrante, D., Pighin, S., & Gonzalez, M. (2007). Postdecisional counterfactual thinking by actors and readers. *Psychological Science, 18*(6), 510-515. doi: 10.1111/j.1467-9280.2007.01931.x

Girotto, V., Legrenzi, P., & Rizzo, A. (1991). Event Controllability in Counterfactual Thinking. *Acta Psychologica, 78*(1-3), 111-133. doi: 10.1016/0001-6918(91)90007-M

Gleicher, F., Kost, K. A., Baker, S. M., Strathman, A. J., Richman, S. A., & Sherman, S. J. (1990). The role of counterfactual thinking in judgments of affect. *Personality and Social Psychology Bulletin, 16*(2), 284-295. doi: 10.1177/0146167290162009

Glisky, E. L. (2007). Changes in cognitive function in human aging. In D. R. Riddle (Ed.), *Brain Aging: Models, Methods, and Mechanisms* (pp. 3-20): CRC Press.

Gneezy, U., Rockenbach, B., & Serra-Garcia, M. (2013). Measuring lying aversion. *Journal of Economic Behavior & Organization, 93*, 293-300. doi: 10.1016/j.jebo.2013.03.025

Golden, C. J., & Freshwater, S. M. (2002). *Stroop Color and Word Test: A Manual for Clinical and Experimental Uses*. Chicago, IL: Stoelting Co.

Golden, C. J., Osmond, D. C., Moses, J. A., & Berg, R. A. (1981). *Interpretation of the Halstead-Reitan Neuropsychological Test Battery*. New York, NY: Grune & Stratton.

Goldman, A. I. (2012). Theory of mind. In E. Margolis, R. Samuels & S. Stich (Eds.), *The Oxford Handbook of Philosophy of Cognitive Science* (pp. 402-424).

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. doi: 10.3200/Mono.132.3.197-214

Goodman, N. (1973). *Fact, Fiction and Forecast. New York*. New York: The Bobbs-Merril Co.

Gotham, A., Brown, R., & Marsden, C. (1988). 'Frontal'cognitive function in patients with Parkinson's disease 'on'and 'off'levodopa. *Brain, 111*(2), 299-321. doi: 10.1093/brain/111.2.299

Green, D. M., & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. New York, NY: Wiley.

Greenberg, A., Smeets, P., & Zhurakhovska, L. (2014). Lying, Guilt, and Shame: Mimeo.

Guajardo, N. R., McNally, L. F., & Wright, A. (2016). Children's spontaneous counterfactuals: The roles of valence, expectancy, and cognitive flexibility. *Journal of Experimental Child Psychology, 146*, 79-94. doi: 10.1016/j.jecp.2016.01.009

Guajardo, N. R., Parker, J., & Turley-Ames, K. J. (2009). Associations among false belief understanding, counterfactual reasoning, and executive function. *British Journal of Developmental Psychology, 27*(3), 681-702. doi: 10.1348/026151008X357886

Guajardo, N. R., & Turley-Ames, K. J. (2004). Preschoolers' generation of different types of counterfactual statements and theory of mind understanding. *Cognitive Development, 19*(1), 53-80. doi: 0.1016/j.cogdev.2003.09.002

Gunia, B. C., Wang, L., Huang, L., Wang, J., & Murnighan, J. K. (2012). Contemplation and conversation: Subtle influences on moral decision making. *Academy of Management Journal, 55*(1), 13-33. doi: 10.5465/amj.2009.0873

El Haj, M., & Antoine, P. (2018). Relationship between Lower Tendency to Deceive in Aging and Inhibitory Compromise. *Gerontology, 64*(1), 67-73.

Hao, L., & Houser, D. (2011). *Honest Lies*. Discussion Paper.

Harris, P. L., German, T., & Mills, P. (1996). Children's use of counterfactual thinking in causal reasoning. *Cognition, 61*(3), 233-259. doi: 10.1016/S0010-0277(96)00715-9

Hilbig, B. E., & Hessler, C. M. (2013). What lies beneath: How the distance between truth and lie drives dishonesty. *Journal of Experimental Social Psychology, 49*(2), 263-266. doi: 10.1016/j.jesp.2012.11.010

Hooker, C., Roese, N. J., & Park, S. (2000). Impoverished counterfactual thinking is associated with schizophrenia. *Psychiatry, 63*(4), 326-335. doi: 10.1521/00332747.2000.11024925

Hughes, R. (1998). Considering the vignette technique and its application to a study of drug injecting and HIV risk and safer behaviour. *Sociology of Health & Illness, 20*(3), 381-400. doi: 10.1111/1467-9566.00107

Hume, E. C., & Smith, A. (2006). University student ethics: the differential explanatory effect of locus of control. *Academy of Educational Leadership Journal, 10*(3), 49.

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. doi: 10.1016/j.neures.2010.11.001

Jankovic, J. (2008). Parkinson's disease: clinical features and diagnosis. *Journal of Neurology, Neurosurgery & Psychiatry, 79*(4), 368-376. doi: 10.1136/jnnp.2007.131045

Kahneman, D., & Miller, D. T. (1986). Norm Theory - comparing reality to its alternatives. *Psychological Review, 93*(2), 136-153. doi: 10.1037/0033-295x.93.2.136

Kahneman, D., & Tversky, A. (1982a). The psychology of preferences. *Scientific American, 246*(1), 160-173.

Kahneman, D., & Tversky, A. (1982b). The simulation heuristic. In D. Kahneman, P. Slovic & A. Tversky (Eds.), *Judgements under uncertainty: Heuristics and biases.* . New York: Cambridge University Press.

Karton, I., & Bachmann, T. (2011). Effect of prefrontal transcranial magnetic stimulation on spontaneous truth-telling. *Behavioural Brain Research, 225*(1), 209-214. doi: 10.1016/j.bbr.2011.07.028

Kasimatis, M., & Wells, G. L. (1995). Individual differences in counterfactual thinking. In N. J. R. M. Olson (Ed.), *What Might Have Been: The Social Psychology of Counterfactual Thinking* (pp. 81-101). Hillsdale, NJ: Lawrence Erlbaum.

Kline, R., Galeotti, F., & Orsini, R. (2014). When foul play seems fair: dishonesty as a response to violations of just deserts. *SRN Electronic Journal* doi: 10.2139/ssrn.2383182

Knappe, M. L., & Comaden, M. G. (1979). Telling it like it isn't and not telling it like it is: Some thoughts on deceptive communication. *Human Communication Research, 5*(3), 270-285. doi: 0.1111/j.1468-2958.1979.tb00640.x

Knobe, J. (2010). Person as scientist, person as moralist. *Behavioral and Brain Sciences, 33*(4), 315-329. doi: 10.1017/S0140525X10000907

Kouchaki, M., & Desai, S. D. (2015). Anxious, threatened, and also unethical: how anxiety makes individuals feel threatened and commit unethical acts. *Journal of Applied Psychology, 100*(2), 360-375. doi: 10.1037/a0037796

Kray, L. J., Galinsky, A. D., & Wong, E. M. (2006). Thinking within the box: The relational processing style elicited by counterfactual mind-sets. *Journal of Personality and Social Psychology, 91*(1), 33-48. doi: 10.1037/0022-3514.91.1.33

Lelieveld, G.-J., Shalvi, S., & Crone, E. A. (2016). Lies that feel honest: Dissociating between incentive and deviance processing when evaluating dishonesty. *Biological Psychology, 117*, 100-107. doi: 10.1016/j.biopsycho.2016.03.009

Lerner, M. J. (1980). *The Belief in a Just World*. Ontario, Canada: Springer.

Levine, T. R., Anders, L. N., Banas, J., Baum, K. L., Endo, K., Hu, A. D., & Wong, N. C. (2000). Norms, expectations, and deception: A norm violation model of veracity judgments. *Communications Monographs, 67*(2), 123-137. doi: 10.1080/03637750009376500

Levine, T. R., & McCornack, S. A. (2014). Theorizing About Deception. *Journal of Language and Social Psychology, 33*(4), 431-440. doi: 10.1177/0261927x14536397

Lewis, D. (1973). Counterfactuals and comparative possibility. *Journal of Philosophical Logic, 2*(4), 418-446.

Lewis, S. J., Slabosz, A., Robbins, T. W., Barker, R. A., & Owen, A. M. (2005). Dopaminergic basis for deficits in working memory but not attentional set-shifting in Parkinson's disease. *Neuropsychologia, 43*(6), 823-832. doi: 10.1016/j.neuropsychologia.2004.10.001

Litvan, I., Aarsland, D., Adler, C. H., Goldman, J. G., Kulisevsky, J., Mollenhauer, B., . . . Weintraub, D. (2011). MDS task force on mild cognitive impairment in Parkinson's disease: Critical review of PD-MCI. *Movement Disorders, 26*(10), 1814-1824. doi: 10.1002/mds.2382

Lundberg, C. G., & Frost, D. E. (1992). Counterfactuals in Financial Decision-Making. *Acta Psychologica, 79*(3), 227-244. doi: 10.1016/0001-6918(92)90059-M

Malone, B. E., Adams, R. B., Anderson, D. E., Ansfield, M. E., & DePaulo, B. M. (1997). *Strategies of deception and thier correlates over the course of friendship.* Paper presented at the American Psychological Society, Washington D.C.

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. doi: 10.1016/j.bbr.2010.03.024

Markman, K. D., Gavanski, I., Sherman, S. J., & Mcmullen, M. N. (1993). The mental simulation of better and worse possible worlds. *Journal of Experimental Social Psychology, 29*(1), 87-109. doi: 10.1006/jesp.1993.1005

Markman, K. D., Lindberg, M. J., Kray, L. J., & Galinsky, A. D. (2007). Implications of counterfactual structure for creative generation and analytical problem solving. *Personality & Social Psychology Bulletin, 33*(3), 312-324. doi: 10.1177/0146167206296106

Markman, K. D., McMullen, M. N., & Elizaga, R. A. (2008). Counterfactual thinking, persistence, and performance: A test of the Reflection and Evaluation Model. *Journal of Experimental Social Psychology, 44*(2), 421-428. doi: 10.1016/j.jesp.2007.01.001

Markman, K. D., Mizoguchi, N., & McMullen, M. N. (2008). “It would have been worse under Saddam:” Implications of counterfactual thinking for beliefs regarding the ethical treatment of prisoners of war. *Journal of Experimental Social Psychology, 44*(3), 650-654. doi: 10.1016/j.jesp.2007.03.005

Masip, J., Garrido, E., & Herrero, C. (2004). Defining deception. *Anales de Psicología, 20*(1), 147-172.

Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research, 45*(6), 633-644. doi: 10.1509/jmkr.45.6.633

McCloy, R., & Byrne, R. M. (2000). Counterfactual thinking about controllable events. *Memory & Cognition, 28*(6), 1071-1078. doi: 10.3758/BF03209355

McCrea, S. M. (2008). Self-handicapping, excuse making, and counterfactual thinking: consequences for self-esteem and future motivation. *Journal of Personality and Social Psychology, 95*(2), 274. doi: 10.1037/0022-3514.95.2.274

McEleney, A., & Byrne, R. M. J. (2006). Spontaneous counterfactual thoughts and causal explanations. *Thinking & Reasoning, 12*(2), 235-255. doi: 10.1080/13546780500317897

McMullen, M. N. (1997). Affective Contrast and Assimilation in Counterfactual Thinking. *Journal of Experimental Social Psychology, 33*, 77-100.

McMullen, M. N., Markman, K. D., & Gavanski, I. (1995). Living in neither the best nor worst of all possible worlds: Antecedents and consequences of upward and downward counterfactual thinking. In N. Roese & J. M. Olson (Eds.), *What might have been: The social psychology of counterfactual thinking*. New York: Taylor and Francis Group.

McNamara, P., Durso, R., Brown, A., & Lynch, A. (2003). Counterfactual cognitive deficit in persons with Parkinson's disease. *Journal of Neurology, Neurosurgery and Psychiatry, 74*(8), 1065-1070. doi: 10.1136/jnnp.74.8.1065

Medvec, V. H., Gilovich, T., & Madey, S. F. (1995). When Less Is More - Counterfactual Thinking and Satisfaction among Olympic Medalists. *Journal of Personality and Social Psychology, 69*(4), 603-610. doi: 10.1037//0022-3514.69.4.603

Mercier, H., Rolison, J. J., Stragà, M., Ferrante, D., Walsh, C. R., & Girotto, V. (2016). Questioning the preparatory function of counterfactual thinking. *Memory & Cognition*, 1-9. doi: 10.3758/s13421-016-0660-5

Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences, 34*(02), 57-74. doi: 10.1017/S0140525X10000968

Meyerslevy, J., & Maheswaran, D. (1992). When Timing Matters - the Influence of Temporal Distance on Consumers Affective and Persuasive Responses. *Journal of Consumer Research, 19*(3), 424-433. doi: 10.1086/209312

Michikyan, M., Subrahmanyam, K., & Dennis, J. (2014). Can you tell who I am? Neuroticism, extraversion, and online self-presentation among young adults. *Computers in Human Behavior, 33*, 179-183. doi: 10.1016/j.chb.2014.01.010

Miller, D. T., Visser, P. S., & Staub, B. D. (2005). How surveillance begets perceptions of dishonesty: the case of the counterfactual sinner. *Journal of Personality and Social Psychology, 89*(2), 117-128. doi: 10.1037/0022-3514.89.2.117

Miller, G. R., & Stiff, J. B. (1993). *Deceptive communication*: Sage Publications, Inc.

Mitchell, R. W. (1986). A framework for discussing deception. In M. N. S. Mitchell R.W. (Ed.), *Deception: Perspectives on human and nonhuman deceit* (pp. 3-40). Albany: State University of New York Press.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49-100. doi: 10.1006/cogp.1999.0734

Möller, T. (2016). *Deception on facebook: relation between personality, deviation from reality and channels of deception.* University of Twente.

Monin, B., & Jordan, A. H. (2009). The dynamic moral self: A social psychological perspective. In D. Narvaez & D. Lapsley (Eds.), *Personality, identity, and character: Explorations in moral psychology* (pp. 341-354): Cambridge University Press.

Moore, C., & Gino, F. (2013). Ethically adrift: How others pull our moral compass from true North, and how we can fix it. *Research in Organizational Behavior, 33*, 53-77.

Ndubuisi, B., & Byrne, R. M. (2013). *Intentionality and choice*. Paper presented at the The Annual Meeting of the Cognitve Science Society, Berlin.

Niedenthal, P. M., Tangney, J. P., & Gavanski, I. (1994). " If only I weren't" versus" If only I hadn't": Distinguishing shame and guilt in conterfactual thinking. *Journal of Personality and Social Psychology, 67*(4), 585-595. doi: 10.1037/0022-3514.67.4.585

Niepel, C., Mustafić, M., Greiff, S., & Roberts, R. D. (2015). The dark side of creativity revisited: Is students' creativity associated with subsequent decreases in their ethical decision making? *Thinking Skills and Creativity, 18*, 43-52. doi: 10.1016/j.tsc.2015.04.005

Nuijten, M. B., Hartgerink, C. H., van Assen, M. A., Epskamp, S., & Wicherts, J. M. (2016). The prevalence of statistical reporting errors in psychology (1985–2013). *Behavior Research Methods, 48*(4), 1205-1226. doi: 10.3758/s13428-015-0664-2

Peer, E., Acquisti, A., & Shalvi, S. (2014). "I cheated, but only a little": Partial confessions to unethical behavior. *Journal of Personality and Social Psychology, 106*(2), 202. doi: 10.1037/a0035392

Penner, I. K., Kobel, M., Stöcklin, M., Weber, P., Opwis, K., & Calabrese, P. (2012). The Stroop task: comparison between the original paradigm and computerized versions in children and adults. *The Clinical Neuropsychologist, 26*(7), 1142-1153. doi: 10.1080/13854046.2012.713513

Perner, J., Sprung, M., & Steinkogler, B. (2004). Counterfactual conditionals and false belief: A developmental dissociation. *Cognitive Development, 19*(2), 179-201. doi: 10.1016/j.cogdev.2003.12.001

Peskin, J. (1992). Ruse and representations: On children's ability to conceal information. *Developmental Psychology, 28*(1), 84. doi: 10.1037/0012-1649.28.1.84

Pettit, L., McCarthy, M., Davenport, R., & Abrahams, S. (2013). Heterogeneity of letter fluency impairment and executive dysfunction in Parkinson's disease. *Journal of the International Neuropsychological Society, 19*(09), 986-994. doi: 10.1017/S1355617713000829

Pfeiffer, H. C. V., Løkkegaard, A., Zoetmulder, M., Friberg, L., & Werdelin, L. (2014). Cognitive impairment in early-stage non-demented Parkinson's disease patients. *Acta Neurologica Scandinavica, 129*(5), 307-318. doi: 10.1111/ane.12189

Phillips, J., Stelmach, G. E., & Teasdale, N. (1991). What can indices of handwriting quality tell us about Parkinsonian handwriting? *Human Movement Science, 10*(2), 301-314. doi: 10.1016/0167-9457(91)90009-M

Pighin, S., Byrne, R. M. J., Ferrante, D., Gonzalez, M., & Girotto, V. (2011). Counterfactual thoughts about experienced, observed, and narrated events. *Thinking & Reasoning, 17*(2), 197-211. doi: 10.1080/13546783.2011.561598

Priori, A., Mameli, F., Cogiamanian, F., Marceglia, S., Tiriticco, M., Mrakic-Sposta, S., . . . Sartori, G. (2008). Lie-specific involvement of dorsolateral prefrontal cortex in deception. *Cerebral Cortex, 18*(2), 451-455. doi: 10.1093/cercor/bhm088

Prokopcáková, A., & Ruiselová, Z. (2008). Counterfactual thinking as related to anxiety and self-esteem. *Studia Psychologica, 50*(4), 429.

Rafetseder, E., Schwitalla, M., & Perner, J. (2013). Counterfactual reasoning: From childhood to adulthood. *Journal of Experimental Child Psychology, 114*(3), 389-404. doi: 10.1016/j.jecp.2012.10.010

Rasga, C., Quelhas, A. C., & Byrne, R. M. (2016). Children’s reasoning about other’s intentions: False-belief and counterfactual conditional inferences. *Cognitive Development, 40*, 46-59. doi: 10.1016/j.cogdev.2016.08.007

Reitan, R. M. (1958). Validity of the Trail Making Test as an indicator of organic brain damage. *Perceptual and Motor Skills, 8*(3), 271-276. doi: 10.2466/pms.1958.8.3.271

Riggio, R. E., Tucker, J., & Throckmorton, B. (1987). Social skills and deception ability. *Personality and Social Psychology Bulletin, 13*(4), 568-577. doi: 10.1177/0146167287134013

Riggs, K. J., Peterson, D. M., Robinson, E. J., & Mitchell, P. (1998). Are errors in false belief tasks symptomatic of a broader difficulty with counterfactuality? *Cognitive Development, 13*(1), 73-90. doi: 10.1016/S0885-2014(98)90021-1

Roese, N. (1999). Counterfactual thinking and decision making. *Psychonomic Bulletin & Review, 6*(4), 570-578. doi: 10.3758/Bf03212965

Roese, N. J. (1994). The Functional Basis of Counterfactual Thinking. *Journal of Personality and Social Psychology, 66*(5), 805-818. doi: 10.1037//0022-3514.66.5.805

Roese, N. J. (1997). Counterfactual thinking. *Psychological Bulletin, 121*(1), 133-148. doi: 10.1037/0033-2909.121.1.133

Roese, N. J., Black, M., & Olson, J. M. (1994). *Perceived similarity and defensive counterfactual thinking: Evidence for motivational determinants of counterfactual content.* .

Roese, N. J., & Epstude, K. (2017). The Functional Theory of Counterfactual Thinking: New Evidence, New Challenges, New Insights. *Advances in Experimental Social Psychology*. doi: 10.1016/bs.aesp.2017.02.001

Roese, N. J., & Hur, T. (1997). Affective determinants of counterfactual thinking. *Social Cognition, 15*(4), 274-290. doi: 10.1521/soco.1997.15.4.274

Roese, N. J., & Morrison, M. (2009). The Psychology of Counterfactual Thinking. *Historical Social Research-Historische Sozialforschung, 34*(2), 16-26.

Roese, N. J., & Olson, J. M. (1993). The structure of counterfactual thought. *Personality and Social Psychology Bulletin, 19*(3), 312-319. doi: 10.1177/0146167293193008

Roese, N. J., & Olson, J. M. (1995). Outcome Controllability and Counterfactual Thinking. *Personality and Social Psychology Bulletin, 21*(6), 620-628. doi: 10.1177/0146167295216008

Roese, N. J., & Olson, J. M. (1997). Counterfactual thinking: The intersection of affect and function. *Advances in Experimental Social Psychology, Vol 29, 29*, 1-59. doi: 10.1016/S0065-2601(08)60015-5

Roese, N. J., Smallman, R., & Epstude, K. (2017). Do episodic counterfactual thoughts focus on controllable action?: The role of self-initiation. *Journal of Experimental Social Psychology, 73*, 14-23. doi: 10.1016/j.jesp.2017.05.006

Sanna, L. J., & Turley, K. J. (1996). Antecedents to spontaneous counterfactual thinking: Effects of expectancy violation and outcome valence. *Personality and Social Psychology Bulletin, 22*(9), 906-919. doi: 10.1177/0146167296229005

Sarzyńska, J., Falkiewicz, M., Riegel, M., Babula, J., Margulies, D. S., Nęcka, E., . . . Szatkowska, I. (2017). More intelligent extraverts are more likely to deceive. *PloS One, 12*(4), e0176591. doi: 10.1371/journal.pone.0176591

Schacter, D. L., Benoit, R. G., De Brigard, F., & Szpunar, K. K. (2013). Episodic future thinking and episodic counterfactual thinking: Intersections between memory and decisions. *Neurobiology of Learning and Memory*. doi: 10.1016/j.nlm.2013.12.008

Schwartz, B., Ward, A., Monterosso, J., Lyubomirsky, S., White, K., & Lehman, D. R. (2002). Maximizing versus satisficing: happiness is a matter of choice. *Journal of Personality and Social Psychology, 83*(5), 1178. doi: 10.1037/0022-3514.83.5.1178

Schweitzer, M. E., & Hsee, C. K. (2002). Stretching the truth: Elastic justification and motivated communication of uncertain information. *Journal of Risk and Uncertainty, 25*(2), 185-201. doi: 10.1023/A:1020647814263

Serota, K. B., Levine, T. R., & Boster, F. J. (2010). The Prevalence of Lying in America: Three Studies of Self-Reported Lies. *Human Communication Research, 36*(1), 2-25. doi: 10.1111/j.1468-2958.2009.01366.x

Shalvi, S., Dana, J., Handgraaf, M. J., & De Dreu, C. K. (2011). Justified ethicality: Observing desired counterfactuals modifies ethical perceptions and behavior. *Organizational Behavior and Human Decision Processes, 115*(2), 181-190. doi: 10.1016/j.obhdp.2011.02.001

Shalvi, S., Eldar, O., & Bereby-Meyer, Y. (2012). Honesty requires time (and lack of justifications). *Psychological Science, 23*(10), 1264-1270. doi: 10.1177/0956797612443835

Sisco, S. M., Slonena, E., Okun, M. S., Bowers, D., & Price, C. C. (2016). Parkinson's disease and the Stroop color word test: processing speed and interference algorithms. *The Clinical Neuropsychologist, 30*(7), 1104-1117. doi: 10.1080/13854046.2016.1188989

Smallman, R., & Roese, N. J. (2009). Counterfactual thinking facilitates behavioral intentions. *Journal of Experimental Social Psychology, 45*(4), 845-852. doi: 10.1016/j.jesp.2009.03.002

Solca, F., Poletti, B., Zago, S., Crespi, C., Sassone, F., Lafronza, A., . . . Ciammola, A. (2015). Counterfactual thinking deficit in Huntington's disease. *PloS One, 10*(6), e0126773. doi: 10.1371/journal.pone.0126773

Spence, S. A., Hunter, M. D., Farrow, T., Green, R. D., Leung, D. H., Hughes, C. J., & Ganesan, V. (2004). A cognitive neurobiological account of deception: evidence from functional neuroimaging. *Philos Trans R Soc Lond B Biol Sci, 359*(1451), 1755-1762.

Spence, S. A., Kaylor-Hughes, C., Farrow, T. F., & Wilkinson, I. D. (2008). Speaking of secrets and lies: the contribution of ventrolateral prefrontal cortex to vocal deception. *Neuroimage, 40*(3), 1411-1418. doi: 10.1016/j.neuroimage.2008.01.035

Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine, 166*(10), 1092-1097. doi: 10.1001/archinte.166.10.1092

Sporer, S. L., & Schwandt, B. (2007). Moderators of nonverbal indicators of deception: A meta-analytic synthesis. *Psychology, Public Policy and Law, 13*(1), 1-34.

Spreng, R. N., Mar, R. A., & Kim, A. S. (2009). The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. *Journal of Cognitive Neuroscience, 21*(3), 489-510. doi: 10.1162/jocn.2008.21029

Stokke, A. (2013). Lying, Deceiving, and Misleading. *Philosophy Compass, 8*(4), 348-359. doi: 10.1111/phc3.12022

Strauss, E., Sherman, E. M., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary*: American Chemical Society.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*(6), 643. doi: 10.1037/0096-3445.121.1.15

Stuss, D. T., Gallup, G. G., & Alexander, M. P. (2001). The frontal lobes are necessary for theory of mind. *Brain, 124*(2), 279-286.

Sweeny, K., & Vohs, K. D. (2012). On near misses and completed tasks the nature of relief. *Psychological Science, 23*(5), 464-468. doi: 10.1177/0956797611434590

Tabachnick, B. G., & Fidell, L. S. (2001). Using multivariate statistics.

Tal-Or, N., Boninger, D. S., Poran, A., & Gleicher, F. (2004). Counterfactual thinking as a mechanism in narrative persuasion. *Human Communication Research, 30*(3), 301-328. doi: 10.1111/j.1468-2958.2004.tb00734.x

Talwar, V., & Lee, K. (2008). Social and cognitive correlates of children's lying behavior. *Child Development, 79*(4), 866-881. doi: 10.1111/j.1467-8624.2008.01164.x

Teigen, K. H., & Jensen, T. K. (2010). Unlucky Victims or Lucky Survivors? *European Psychologist, 16*, 48-57. doi: 10.1027/1016-9040/a000033

ten Brinke, L., & Porter, S. (2012). Cry me a river: Identifying the behavioral consequences of extremely high-stakes interpersonal deception. *Law and Human Behavior, 36*(6), 469. doi: 10.1037/h0093929

Tracy, J. L., & Robins, R. W. (2006). Appraisal antecedents of shame and guilt: support for a theoretical model. *Personality and Social Psychology Bulletin, 32*(10), 1339-1351. doi: 10.1177/0146167206290212

Trevino, L. K. (1986). Ethical decision making in organizations: A person-situation interactionist model. *Academy of Management Review, 11*(3), 601-617. doi: 10.5465/AMR.1986.4306235

Tyser, M. P., McCrea, S. M., & Knüpfer, K. (2012). Pursuing perfection or pursuing protection? Self-evaluation motives moderate the behavioral consequences of counterfactual thoughts. *European Journal of Social Psychology, 42*(3), 372-382. doi: 10.1002/ejsp.1864

Van Bockstaele, B., Wilhelm, C., Meijer, E., Debey, E., & Verschuere, B. (2015). When deception becomes easy: the effects of task switching and goal neglect on the truth proportion effect. *Frontiers in Psychology, 6*, 1666. doi: 10.3389/fpsyg.2015.01666

Van Der Zee, S., Anderson, R., & Poppe, R. (2016). When lying feels the right thing to do. *Frontiers in Psychology, 7*(734). doi: 10.3389/fpsyg.2016.00734

Van Hoeck, N., Begtas, E., Steen, J., Kestemont, J., Vandekerckhove, M., & Van Overwalle, F. (2014). False belief and counterfactual reasoning in a social environment. *Neuroimage, 90*(15), 315-325. doi: 10.1016/j.neuroimage.2013.12.043

Van Hoeck, N., Watson, P. D., & Barbey, A. K. (2015). Cognitive neuroscience of human counterfactual reasoning. *Frontiers in Human Neuroscience, 420*(9). doi: 10.3389/fnhum.2015.00420

Vendemia, J. M. C., Schillaci, M. J., Buzan, R. F., Green, E. P., & Meek, S. W., . (2009). Alternate technologies for the detection of deception. In D. T. Wilcox (Ed.), *The use of polygraph in the assessing, treating and supervising sex offenders.* (pp. 267-296). West Sussex, UK: Wiley-Blackwell.

Verschuere, B., & De Houwer, J. (2011). Detecting concealed information in less than a second: response latency-based measures. In B. Verschuere, B. S. Gershan & E. Meijer (Eds.), *Memory detection: Theory and application of the Concealed Information Test* (pp. 46-62). Cambridge: Cambridge University Press.

Vrij, A. (2000). *Detecting lies and deceit: The psychology of lying and implications for professional practice*. Chichester: Wiley.

Vrij, A. (2007). Deception: A social lubricant and a selfish act. In K. Fiedler (Ed.), *Social communication* (pp. 309-342). Hove: Psychology Press.

Vrij, A., Fisher, R., Mann, S., & Leal, S. (2008). A cognitive load approach to lie detection. *Journal of Investigative Psychology and Offender Profiling, 5*(1-2), 39-43. doi: 10.1002/jip.82

Vrij, A., & Granhag, P. A. (2012). Eliciting cues to deception and truth: What matters are the questions asked. *Journal of Applied Research in Memory and Cognition, 1*(2), 110-117. doi: 10.1016/j.jarmac.2012.02.004

Vrij, A., Granhag, P. A., & Mann, S. (2010). Good liars. *Journal of Psychiatry and Law, 38*(1-2), 77-98. doi: 10.1177/009318531003800105

Vrij, A., Leal, S., Mann, S. A., & Granhag, P. A. (2011). A comparison between lying about intentions and past activities: Verbal cues and detection accuracy. *Applied Cognitive Psychology, 25*(2), 212-218. doi: 10.1002/acp.1665

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. doi: 10.1002/1099-0720(200103/04)15:23.0.CO;2-A

Vrij, A., Mann, S., Kristen, S., & Fisher, R. P. (2007). Cues to deception and ability to detect lies as a function of police interview styles. *Law and Human Behavior, 31*(5), 499-518. doi: 10.1007/s10979-006-9066-4

Vrij, A., Semin, G. R., & Bull, R. (1996). Insight into behavior displayed during deception. *Human Communication Research, 22*(4), 544-562.

Walczyk, J. J., Harris, L. L., Duck, T. K., & Mulay, D. (2014). A social-cognitive framework for understanding serious lies: Activation-decision-construction-action theory. *New Ideas in Psychology, 34*, 22-36. doi: 10.1016/j.newideapsych.2014.03.001

Walczyk, J. J., Mahoney, K. T., Doverspike, D., & Griffith-Ross, D. A. (2009). Cognitive lie detection: response time and consistency of answers as cues to deception. *Journal of Business and Psychology, 24*(1), 33-49. doi: 10.1007/s10869-009-9090-8

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. doi: 10.1002/Acp.914

Walczyk, J. J., Schwartz, J. P., Clifton, R., Adams, B., Wei, M., & Zha, P. (2005). Lying person-to-person about life events: A cognitive framework for lie detection. *Personnel Psychology, 58*(1), 141-170. doi: 10.1111/j.1744-6570.2005.00484.x

Walsh, C. R., Deeprose, C., & Briazu, R. A. (under review). *Counterfactual thinking and aging*.

Warden, C., Hwang, J., Marshall, A., Fenesy, M., & Poston, K. L. (2016). The effects of dopamine on digit span in Parkinson's disease. *Journal of Clinical Movement Disorders, 3*(1), 5. doi: 10.1186/s40734-016-0033-z

Watkins, E. R. (2008). Constructive and unconstructive repetitive thought. *Psychological Bulletin, 134*(2), 163-206. doi: 10.1037/0033-2909.134.2.163

Wechsler, D. (1997). *Wechsler memory scale (WMS-III)*: Psychological Corporation San Antonio, TX.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child Development, 72*(3), 655-684. doi: 10.1111/1467-8624.00304

Wells, G. L., Taylor, B. R., & Turtle, J. W. (1987). The Undoing of Scenarios. *Journal of Personality and Social Psychology, 53*(3), 421-430. doi: 10.1037/0022-3514.53.3.421

West, C., & Zhong, C.-B. (2015). Moral cleansing. *Current Opinion in Psychology, 6*, 221-225. doi: 10.1016/j.copsyc.2015.09.022

Williams, E. J., Bott, L. A., Patrick, J., & Lewis, M. B. (2013). Telling Lies: The Irrepressible Truth? *Plos One, 8*(4). doi: 10.1371/journal.pone.0060713

Wilson, D. S., & O'Gorman, R. (2003). Emotions and actions associated with norm-breaking events. *Human Nature, 14*(3), 277-304. doi: 10.1007/s12110-003-1007-z

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

Xu, H., Bègue, L., & Bushman, B. J. (2012). Too fatigued to care: Ego depletion, guilt, and prosocial behavior. *Journal of Experimental Social Psychology, 48*(5), 1183-1186. doi: 10.1016/j.jesp.2012.03.007

Yang, Y., Raine, A., Lencz, T., Bihrle, S., Lacasse, L., & Colletti, P. (2005). Prefrontal white matter in pathological liars. *The British Journal of Psychiatry, 187*(4), 320-325. doi: 10.1016/j.jesp.2012.03.007

Yokota, S., Taki, Y., Hashizume, H., Sassa, Y., Thyreau, B., Tanaka, M., & Kawashima, R. (2013). Neural correlates of deception in social contexts in normally developing children. *Frontiers in Human Neuroscience, 7*. doi: 10.3389/fnhum.2013.00206

Zeelenberg, M., van Dijk, W. W., van der Pligt, J., Manstead, A. S. R., van Empelen, P., & Reinderman, D. (1998). Emotional reactions to the outcomes of decisions: The role of counterfactual thought in the experience of regret and disappointment. *Organizational Behavior and Human Decision Processes, 75*(2), 117-141. doi: 10.1006/obhd.1998.2784

Zelazo, P. D., Müller, U., Frye, D., Marcovitch, S., Argitis, G., Boseovski, J., . . . Sutherland, A. (2003). The development of executive function in early childhood. *Monographs of the society for research in child development*, i-151.

Zimbler, M., & Feldman, R. S. (2011). Liar, liar, hard drive on fire: How media context affects lying behavior. *Journal of Applied Social Psychology, 41*(10), 2492-2507. doi: 10.1111/j.1559-1816.2011.00827.x

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 14, pp. 1-59). New York: Academic Press.