THE UNIVERSITY OF WARWICK

Library Declaration and Deposit Agreement

1. STUDENT DETAILS

Please complete the following:

Full name: ……

University ID number: ……

2. THESIS DEPOSIT

2.1 I understand that under my registration at the University, I am required to deposit my thesis with the University in BOTH hard copy and in digital format. The digital version should normally be saved as a single pdf file.

2.2 The hard copy will be housed in the University Library. The digital version will be deposited in the University's Institutional Repository (WRAP). Unless otherwise indicated (see 2.3 below) this will be made openly accessible on the Internet and will be supplied to the British Library to be made available online via its Electronic Theses Online Service (EThOS) service.
[At present, theses submitted for a Master's degree by Research (MA, MSc, LLM, MS or MMedSci) are not being deposited in WRAP and not being made available via EthOS. This may change in future.]

2.3 In exceptional circumstances, the Chair of the Board of Graduate Studies may grant permission for an embargo to be placed on public access to the hard copy thesis for a limited period. It is also possible to apply separately for an embargo on the digital version. (Further information is available in the *Guide to Examinations for Higher Degrees by Research*.)

2.4 *If you are depositing a thesis for a Master's degree by Research, please complete section (a) below. For all other research degrees, please complete both sections (a) and (b) below:*

(a) Hard Copy

I hereby deposit a hard copy of my thesis in the University Library to be made publicly available to readers (please delete as appropriate) EITHER immediately OR after an embargo period of ……..... months/years as agreed by the Chair of the Board of Graduate Studies.

I agree that my thesis may be photocopied. YES / NO (*Please delete as appropriate*)

(b) Digital Copy

I hereby deposit a digital copy of my thesis to be held in WRAP and made available via EThOS.

Please choose one of the following options:

EITHER My thesis can be made publicly available online. YES / NO (*Please delete as appropriate*)

OR My thesis can be made publicly available only after…..[date] (Please give date)
YES / NO (*Please delete as appropriate*)

OR My full thesis cannot be made publicly available online but I am submitting a separately identified additional, abridged version that can be made available online.
YES / NO (*Please delete as appropriate*)

OR My thesis cannot be made publicly available online. YES / NO (*Please delete as appropriate*)

3. GRANTING OF NON-EXCLUSIVE RIGHTS

Whether I deposit my Work personally or through an assistant or other agent, I agree to the following:

Rights granted to the University of Warwick and the British Library and the user of the thesis through this agreement are non-exclusive. I retain all rights in the thesis in its present version or future versions. I agree that the institutional repository administrators and the British Library or their agents may, without changing content, digitise and migrate the thesis to any medium or format for the purpose of future preservation and accessibility.

4. DECLARATIONS

(a) I DECLARE THAT:

- I am the author and owner of the copyright in the thesis and/or I have the authority of the authors and owners of the copyright in the thesis to make this agreement. Reproduction of any part of this thesis for teaching or in academic or other forms of publication is subject to the normal limitations on the use of copyrighted materials and to the proper and full acknowledgement of its source.

- The digital version of the thesis I am supplying is the same version as the final, hard-bound copy submitted in completion of my degree, once any minor corrections have been completed.

- I have exercised reasonable care to ensure that the thesis is original, and does not to the best of my knowledge break any UK law or other Intellectual Property Right, or contain any confidential material.

- I understand that, through the medium of the Internet, files will be available to automated agents, and may be searched and copied by, for example, text mining and plagiarism detection software.

(b) IF I HAVE AGREED (in Section 2 above) TO MAKE MY THESIS PUBLICLY AVAILABLE DIGITALLY, I ALSO DECLARE THAT:

- I grant the University of Warwick and the British Library a licence to make available on the Internet the thesis in digitised format through the Institutional Repository and through the British Library via the EThOS service.

- If my thesis does include any substantial subsidiary material owned by third-party copyright holders, I have sought and obtained permission to include it in any version of my thesis available in digital format and that this permission encompasses the rights that I have granted to the University of Warwick and to the British Library.

5. LEGAL INFRINGEMENTS

I understand that neither the University of Warwick nor the British Library have any obligation to take legal action on behalf of myself, or other rights holders, in the event of infringement of intellectual property rights, breach of contract or of any other right, in the thesis.

Please sign this agreement and return it to the Graduate School Office when you submit your thesis.

Student's signature: .....…… Date: .....

The mediating and moderating factors of fabricated evidence on false confessions, beliefs and memory

By

Deborah S. Wright

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy in Psychology

University of Warwick, Department of Psychology

April 2013

Contents

List of Tables

List of Figures

Acknowledgments

I would like to start by thanking my supervisor, Kim, for inspiring me and providing me with great opportunities. Thanks for your insanely speedy email replies and theoretical brilliance! I would also like to thank my secondary supervisor, Derrick, for his faith in me, and for not giving up on the driving task method! And of course, my mentor, Neil, for his immense knowledge of all things Psychology! Thanks also to all the RAs who volunteered to help me over the three years, especially the enthusiastic and ever cheerful Kay Chiang.

A big shout out goes to my fellow postgraduate students. Most notably, my partner in crime, Lucinda, and stats guru, Anna, who have both become firm friends. Indeed, thanks to all the *Center Parcs* crew for supporting me and providing an epic amount of light relief!

A special mention goes to Tom, for motivating me, keeping me calm, and never wavering in his enthusiasm for my research and belief in me!

Finally, the biggest thanks of all has to go to my family, Vic's, Mum, Dad and Nan. Dad sadly passed away before my graduation... So this is especially for you, Dad. Thank you all for being there to support me throughout, and for providing me with the strength and determination to see it through. And most of all, for being proud of me and letting everyone know it!

Declaration

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree. The work presented (including data generated and data analysis) was carried out by the author. However, some parts of this thesis have been published in an academic journal. Contributions are as follows:

Experiments 1 and 2/Chapters 4 and 5:

Wright, D. S., Wade, K. A., & Watson, D. G. (2013). Delay and déjà vu: Timing and repetition increase the power of false evidence. *Psychonomic Bulletin & Review*. doi: 10.3758/s13423-013-0398-z

Idea: Wright, Wade and Watson

First draft: Wright

Statistical Analysis: Wright

Subsequent revisions: Wright, Wade and Watson

One ancillary paper has also been submitted for publication:

Wright, D. S., Nash, R. A., & Wade, K. A. (Submitted, January 2013). Good cop, bad cop: Rapport-building induces eyewitnesses to corroborate false evidence. *Law and Human Behavior*

Deborah S. Wright

Abstract

Fabricated evidence (e.g., doctored videos) can induce people to falsely confess to committing a 'crime' and change the way people remember an event. This happens because memory is both malleable and reconstructive: people remember their past using information available to them in the present. Sometimes people rely on external evidence to tell them what happened, which may or may not be accurate.

A wide range of studies have demonstrated the robust and persuasive effect of false evidence in different situations, and some have examined the theoretical mechanisms behind the effect. However, little is known about factors that might mediate or moderate the power of false evidence. This thesis examines some of these factors and their behavioural consequences.

Experiments 1 and 2 use a novel method to investigate whether changes in when false evidence is presented, or how many times it is presented, make false evidence more powerful. The results highlight the importance of several theoretical mechanisms, which are then explored in the succeeding chapters.

Experiments 3 and 4 use the same method to examine whether the type of evidence presented, or the order in which it is presented, influences its effect. The findings build upon those of Experiments 1 and 2 and suggest some interesting interactions between the different moderating and mediating factors.

Combined with a questionnaire examining peoples' perceptions of digitally edited materials, the findings of Experiments 1-4 have practical and methodological implications. Importantly, the results also have potential theoretical implications and suggest modifications to Mazzoni and Kirsch's (2002) model of autobiographical belief and memory. The amended model includes the role of the examined moderating and mediating factors, and thus is better able to account for how external evidence influences memory processes. In sum, this thesis helps us to understand situations in which false evidence might be particularly powerful.

Chapter 1: False confessions in the real world and the laboratory

"There is a luxury in self-reproach. When we blame ourselves, we feel that no one else has a right to blame us. It is the confession, not the priest, that gives us absolution."

—Oscar Wilde (1854-1900)

Most people do not think they would ever confess to doing something they did not do. The idea in itself seems counterintuitive and people find it hard to understand the reasons behind a false confession. This is one of the main problems that false confessors face in court and it could be the driving force behind their high conviction rates: people simply do not believe their confession was false because they do not think they would ever falsely confess themselves. This thesis examines factors that contribute to people making false confessions, and more specifically what happens when people are presented with fabricated evidence that implicates them in a crime. In the first part of this chapter, the types, prevalence, reasons behind, and consequences of false confessions are discussed. The second part of this chapter then outlines how false confessions have been investigated through experiments.

Part I: False confessions in the real world

What are false confessions?

A person falsely confesses when they admit to committing a crime they did not actually commit. Kassin and Wrightsman (1985) classified false confessions into three types. This taxonomy has since been revised (e.g., Gudjonsson, 2003) but remains the simplest and arguably the most useful taxonomy of false confessions. First, there are *voluntary* confessions, in which a person admits to doing something without any pressure from the police. This can occur if the person has a desire for infamy, wants to punish themselves, offload their guilt, or protect someone else, or if they have difficulty distinguishing fact and fiction. Second, there are *coerced-compliant* confessions, in which a person confesses due to police pressure. During interrogation, the immediate gain—for instance, of being able to go home or avoid further interrogation—is perceived to outweigh the long-term consequences, such as being convicted. Third, there are *coerced-internalised* confessions, in which a person comes to erroneously believe they committed the crime, and sometimes even develops false memories of doing it. This can happen if the person has reason to doubt their own memory and therefore relies on external cues to make sense of the situation (Gudjonsson, 2003).

Sometimes, manipulative police questioning can lead to a coerced-compliant false confession becoming a coerced-internalised one. This is suggested in the case of Carole Richardson of the Guildford Four (Gudjonsson, 2003). Along with three others, she was convicted and sentenced to life imprisonment for the IRA bombings of two public houses in Surrey, UK, in 1974. She initially made a coerced confession during interrogation after constant questioning and alleged physical violence by police officers. This confession became internalised

for several days, which was exacerbated by her confusion over officers' unwavering confidence that she had committed the crime. Richardson was finally released with the other three suspects in 1989 after 15 years in prison, when the Court of Appeals quashed her conviction.

How often do false confessions occur?

False confessions have been implicated in 20-25% of cases in the US, in which the suspect has later been exonerated through DNA evidence proving them innocent (Garrett, 2008; http://www.innocenceproject.org; Kassin, Bogart, & Kerner, 2012). It is estimated that in reality the figures are much higher, especially for capital murder cases (Drizin & Leo, 2004; White, 2003). However, it is almost impossible to estimate how often false confessions occur because it is hard to prove that someone is innocent and records are often incomplete, inaccessible, or missing (Gudjonsson, 2003; Leo, 2008).

Surveys conducted in the US vary in their prevalence estimates. In a survey of over 150 law enforcement officers, 25% of the sample believed they had taken a false confession, with the majority saying that they knew because the confession did not match other evidence (Russano & Narchet, 2009). However, another survey across the US found that police investigators believed that innocent suspects provide a false confession on average only 5% of the time compared to 69% of guilty suspects (Kassin et al., 2007). Respondents also reported having seen an average of 0.7 false confessions themselves. Ramsey and Frank (2007) surveyed almost 800 Ohio criminal justice professionals including

police, prosecutors, defence lawyers and judges. Wrongful conviction for serious cases was perceived to occur in around 1-3% of US cases.

Interestingly, the general public perceive false confession rates to be much higher than criminal justice professionals. For instance, in a survey of over 500 jury-eligible US citizens from 38 states, 67% thought an innocent person accused of a crime would confess after strenuous pressure, and 52% thought a jury would find an innocent confessor guilty (Chojnacki, Cicchini, & White, 2008). In another study survey, respondents estimated that 25% of people who are arrested falsely confess to a crime (Henkel, Coffman, & Dailey, 2008). Paradoxically, Henkel and colleagues' sample held the belief that although other people may falsely confess to a crime, they themselves would be unlikely to confess.

In the UK, the *Police and Criminal Evidence* (*PACE*) *Act* was introduced in 1984 in an attempt to reform policing practice and combat wrongful conviction. Shawyer, Milne and Bull (2009) describe the turn of events in their paper, *Investigative interviewing in the UK.* The PACE act was introduced after several miscarriages of justice in the 1970s, and an observational study highlighting coercive and manipulative tactics being used by the police to force confessions (Irving & Hilgendorf, 1980). Along with further reforms—such as the introduction of an ethical interviewing framework—*interrogation* changed to *investigative interviewing*[1]with the aim of information gathering as opposed to obtaining a confession. The emphasis is now on truth seeking in a fair way with an open mind, in order to reduce the risk of false confessions. This is done by taking away the coercion and pressure and making police aware of more

[1] In this thesis, "interview" is used specifically to refer to UK practices only and "interrogation" is used in all other situations.

vulnerable and suggestible adults so they can prepare for the interview with this in mind. Most studies show that although there are still improvements to be made, the developments have been largely successful, with police detectives perceiving a change from a "confession culture" to interviewing with the aim of seeking the truth (Soukara, Bull, & Vrij, 2002).

How might police tactics induce false confessions?

Psychological scientists around the world have been studying the various factors that could induce or contribute to the problem of false confessions. Much of this research has focussed on the interrogation tactics currently used by law enforcement officers, particularly officers in the US. Previously, interrogation focussed on using *third degree* methods—such as causing physical pain—to obtain a confession. However, Inbau and colleagues introduced a new method in 1986 called the *Reid technique.* This technique was laid out in a manual designed for police interrogators, with the focus being on interrogation as a *social process* between the interrogator and suspect (Inbau, Reid, & Buckley, 1986). The *Reid technique* is now used widely across the US, whereas a similar method introduced in the UK proved not nearly as popular (Walkley, 1987). In their manual, Inbau and colleagues (2005) define interrogation as "an accusational interaction with a suspect, conducted in a controlled environment, designed to persuade the suspect to tell the truth" (p. 3). This definition makes interrogation a "guilt-presumptive process" whereby officers assume a suspect is guilty and interpret all subsequent evidence with this in mind (Leo, 2008). The aim of the process is to persuade the

suspect to overcome their denials and, ultimately, to make a verbal and written confession statement.

The Reid technique consists of nine main steps that the interrogator should follow to obtain a confession. First, the interrogator must directly state their certainty in suspect's guilt. Second, they must identify and develop "themes", which provide excuses or downplay the seriousness of the crime. Third and fourth, they must interrupt and overcome the suspect's denials and objections. Fifth, the interrogator must attempt to hold the suspect's attention by using appropriate body language. Sixth, they must then interpret the suspect's body language to see if they are displaying signs of resignation. Seventh, the interrogator should present the suspect with an alternative motive for the crime, which appears much worse than their actual potential reason. Eighth, the interrogator must obtain an oral confession from the suspect, which includes details about the crime. Finally, the interrogator must obtain a written confession, after the suspect has been reminded about their right to remain silent.

The authors claim that all of these nine steps are legally and morally justifiable. In a later (2005) version of the manual, the authors acknowledge false confessions as a problem, but they attribute false confessions to age, mental impairments and long interrogations using illegal tactics, rather than their recommended techniques. However, according to Professor Richard Leo (2008)—a pioneer of interrogation research and expert witness—using methods like the Reid technique, the police induce false confessions in three ways. First, they make a *misclassification error,* in which they erroneously decide the suspect is guilty. This makes the suspect begin to doubt themselves because of the investigator's

certainty of their guilt and repeated accusations. Second, the police make a *coercion error,* in which they use coercive interrogation techniques to force the suspect to confess. During this time, the suspect is made to feel as if there is no alternative to making a confession. The suspect starts to imagine how they could have committed the crime because the police make it seem plausible. For instance, the investigator may suggest that the suspect "repressed" memories of the terrible event or that the suspect "blacked out", and thus, it is plausible that they might not remember what happened. Third, the police make a *contamination error,* in which they make suggestions to the suspect and provide them with details of the crime that have not been disclosed to the public. The suspect then uses these details about the crime to make up and confabulate further details about how they committed the crime.

Worryingly, surveys show that police officers do not fully understand the psychological reasons for false confessions, or the dangers of using certain interrogation techniques. They still self-report using elements of the Reid technique and tactics such as minimising the blame or seriousness of the crime and claiming false evidence (Kassin et al., 2007; Meyer & Reppucci, 2007), which have been shown to lead innocent people to confess (e.g., Kassin & Kiechel, 1996; Russano, Meissner, Narchet, & Kassin, 2005). Only 7% of a sample of 164 law enforcement officers in a US survey cited poor interrogation techniques as the most common reason for an innocent suspect confessing (Russano & Narchet, 2009). In another survey, 75% of a sample of 99 police department administrators across the US denied that interrogation techniques can sometimes help to produce false confessions (Zalman & Smith, 2007). The

general public on other hand, believe that interrogations involve coercive tactics but that these tactics are necessary in eliciting truthful confessions and are generally not likely to elicit false confessions (Henkel et al., 2008).

How influential is confession evidence?

Confession evidence is so potent that the majority of false confessors are convicted at trial even if they plead not guilty (Drizin & Leo, 2004; Leo & Ofshe, 1998), and this occurs despite a lack of corroborating evidence (Leo, 2008). This is the outcome even when the juries are told the confessor was coerced and stressed during interrogation (Henkel, 2008; Kassin & Sukel, 1997). Confession evidence is powerful enough to "taint" perceptions of other evidence in a case, thus leading to a *bias snowball effect* (Dror, 2012). For instance, if a fingerprint expert knows that a suspect has confessed—and thus concludes that they are likely to be guilty—the expert may then be more likely to interpret a fingerprint match. If more pieces of evidence are then collected and not kept independent, they increasingly influence each other, and the greater the bias becomes. Indeed, research has shown that confession evidence has the potential to influence judgments made by eyewitnesses, jurors, judges, finger-print experts and polygraph examiners (Dror, Charlton, & Peron, 2006; Elaad, Ginton, & Ben-Shakhar, 1994; Hasel & Kassin, 2009; Kassin & Sukel, 1997; Wallace & Kassin, 2012). It also often precedes other errors, such as improper forensic science, mistaken eyewitness identification and informant errors (Kassin et al., 2012).

The power of false confessions over other forms of evidence is most clearly depicted in the case of the "Norfolk Four" (Wells & Leo, 2008). Four men

were wrongfully convicted of the rape and murder of Michelle Bosko in 1998, in Virginia, US. Despite solid alibis, a complete lack of DNA evidence linking the men to the crime scene, and vast inconsistencies between details of the crime in their statements, all four men were sentenced to prison because of coerced false confessions that were later retracted.

Recent research suggests that confession evidence is particularly powerful when it includes specific details about how and why the supposed perpetrator committed the crime (Appleby, Hasel, & Kassin, 2011). In a content analysis of 20 false confessions made by DNA exonerees, all statements included visual details, details about the site and location of the crime, and information about the victim and their behaviour. Often the details given by false confessors are accurate and include details that were not disclosed to the public (Garrett, 2010). As Leo (2008) points out, these "nonpublic" details appear to corroborate the confession and serve to almost guarantee a conviction. Indeed, mock-jurors have been found to be more confident in their guilty verdicts when "how" and "why" details are included in the confession statement (Appleby et al., 2011).

Is evidence disclosed tactically?

Sometimes, the Police not only disclose nonpublic information, but also refer to—and present—evidence. For example, one of the suspects in the Norfolk Four case was shown a photograph of the body of the victim during interrogation. In the UK, the investigator is under no obligation to disclose any evidence against the suspect to the defence solicitor prior to interview (Association of Chief Police Officers (ACPO), 2005). Rather, it is legal to withhold, select and *drip-feed*

evidence over time throughout an interview (Sanders & Young, 2007). The introduction of the British PACE Act (1984) made it illegal for police investigators to lie to suspects about evidence to induce confessions. However, in the US, it is still legal in some jurisdictions—and indeed encouraged by some training manuals (e.g., Inbau et al., 2005)—to make false claims about evidence or actually present suspects with false evidence during interrogation. In the case of the Norfolk Four, suspects were told that an eyewitness had seen them, and one suspect was told there was DNA evidence against him and that he had failed a polygraph test. Interestingly, the general public view the use of false evidence during interrogation as a coercive tactic, yet see it as unlikely to induce false confessions (Leo & Liu, 2009).

In a survey of 40 UK police detectives, disclosure of legitimate evidence was seen as the number one persuasive tactic when interviewing uncooperative suspects (Soukara et al., 2002). Indeed, disclosure of evidence is a common tactic used in all stages of interrogation, and has the potential to influence the suspect confessing (Soukara, Bull, Vrij, Turner, & Cherryman, 2009). One study suggests that evidence disclosure is also commonly used in the US (Leo, 1996). In approximately 90% of 182 suspect interviews for felonies, investigators were found to have confronted the suspect with either true or false evidence of their guilt followed by suggesting that confessing would advance self-interest (Leo, 1996). Most of the time, police officers used initial negative incentives, and then contrasted these with positive incentives. However, anecdotal practitioner literature also suggests that sometimes investigators might use the opposite "Mum's the word" strategy during interrogation (Kalbfleisch, 1994). Here, the

investigator actually minimises disclosure to avoid 'slips' of information and to make the suspect think that the investigator knows more than they do in reality.

Disclosure of evidence to a suspect may not be a problem if the suspect is guilty. Indeed, it may have the intended effect of making the suspect realise that the evidence against them is overwhelming, and thus that they have no other choice but to confess. However, the same unintentional effect may also be true for innocent suspects, who become resigned to their fate. There is also the added problem for innocent suspects, in that the investigator commits a contamination error—as discussed earlier in the chapter—by providing the suspect with information about the crime that the suspect might then later use to confabulate details. Thus, disclosing evidence to an innocent suspect may induce a highly detailed and convincing false confession that ultimately leads to their wrongful conviction.

Part II: False confessions in the laboratory

The first part of this chapter discussed false confessions in the real world and how coercive interrogation techniques might induce innocent people to confess to committing a crime. The second part of this chapter continues the exploration of false confessions by discussing findings from experimental studies. This thesis specifically investigates how and why people make both coerced-compliant and coerced-internalised confessions. Thus the research discussed in this section focuses not only on how people come to comply, but also how they come to falsely believe and remember aspects of events—or indeed entire

events—that never happened. The section starts by outlining early research investigating what happens when people are supplied with erroneous information after an event. Then, research is discussed that investigates how text and images can change belief and memory. Following this, studies are outlined that show how false evidence in particular can alter not only belief and memory, but importantly behaviour too. Finally, the relevance of the plausibility of the overall suggestion is discussed.

The power of misinformation to change memory

Bartlett (1932) famously showed that memory is not necessarily an accurate record of events. Elizabeth Loftus returned to this idea several decades later, when she conducted a series of experiments and became the founder of modern eyewitness testimony research (e.g., Loftus, 2005; Loftus, Miller, & Burns, 1978; Loftus & Palmer, 1974). In these experiments, she demonstrated that information provided after an event—*Postevent Information* (*PEI*)—can alter our memory of that event.

Loftus pioneered the *misinformation method*. The method involves showing participants an event, providing them with misleading information about that event and then testing them to see if they have incorporated this misinformation into memory. For instance, in one of the first studies, participants were shown clips of traffic collisions and afterwards asked how fast the cars were travelling when they ****s*mashed* into each other (Loftus & Palmer, 1974). The **** verb 'smashed' was replaced with one of five words—collided, bumped, hit, contacted—to determine whether a subtle change in wording could result in

dramatic changes in memory. Results showed that the phrasing of the question influenced participants' estimates of speed, with 'smashed' eliciting higher estimates than 'contacted'. In another similar study, Loftus (2005) showed that participants "remembered" seeing a barn, after they were asked how fast a car was travelling in a series of slides when it passed a barn. In reality, no barn featured on the slides, and thus Loftus showed that leading questions can induce people to develop a memory for a nonexistent object. The idea that leading questions can damage eyewitness testimony has become a widely recognised phenomenon, and police officers are now taught to avoid using them during questioning (e.g., Bryant, 2006).

Since the early studies, many more experiments have been conducted to build on the misinformation and leading question findings. For example, simply asking people a leading question about whether or not they have seen video footage of an event, can lead people to say that they viewed the footage, even if the footage does not actually exist. The car crash that killed Princess Diana for instance, was never actually filmed, yet people respond to the leading question by saying not only that they saw the footage, but also by describing the nonexistent footage in detail (e.g., Crombag, Wagenaar, & van Koppen, 1996; Ost, Vrij, Costall, & Bull, 2002). Another study showed that it is much easier to mislead participants about peripheral details than central, more salient information (e.g., Dritsas & Hamilton, 1977, as cited in Loftus, 1979). This finding has since been extended to explain the *weapon focus* effect, in which witnesses who see a crime conducted where a weapon is present, tend to focus centrally on the weapon, and thus do not have a good memory for other event details. This was clearly

demonstrated in an early experimental study, in which participants were more accurate in line-up decisions when they had seen the supposed perpetrator of a mock-crime holding a pen than when he had been holding a bloodied knife (Johnson & Scott, 1976, as cited in Loftus, 1979).

Loftus (1977) also showed that participants' memories can become a compromise between the original event and PEI. In a similar experiment to Loftus and Palmer (1974), participants viewed a green car in an original slide, but were told it was blue afterwards. Participants then reported a compromise memory of having seen a bluish-green car. This idea has since been extended using the MORI technique, in which witnesses simultaneously view what they think is the same video of an event, and are then asked to discuss the event and answer questions about it. In reality, the witnesses see differing versions of the video because they wear special glasses that alter certain details (e.g., the colour of the car). Studies using this technique show that witnesses often incorporate details from each others' accounts of an event, and in doing so can falsely report another person of being guilty of a crime (e.g., Gabbert, Memon, & Allan, 2003; Garry, French, Kinzett, & Mori, 2008). This might occur because participants mistakenly think that the information the other person gave them actually came from the event itself, and thus they misreport it. Indeed, studies have shown that when warned and asked to recall the source of the information, the majority of participants erroneously state that they remember the details from the original event rather than from PEI provided by a confederate (Meade & Roediger, 2002; Roediger, Meade, & Bergman, 2001). This suggests that participants are not simply conforming, rather that their memories are being altered in some way. Thus,

witnesses should always be asked if they have discussed the event in question with other co-witnesses, and if they have, their testimony and recall of events should not be considered independent (Wright, Memon, Skagerberg, & Gabbert, 2009).

Loftus originally argued that the misinformation method demonstrated the malleability of human memory, such that the PEI overwrites, impairs and permanently alters the event memory. However, other theorists argued that the original event memory remains intact but the PEI makes it harder to access (e.g., Bowers & Bekerian, 1984; Christiaansen & Ochalek, 1983). Yet other theorists claimed that neither impairment nor inaccessibility were viable mechanisms, rather that participants are simply biased to accept misinformation (e.g., McCloskey & Zaragoza, 1985). Because the experiments rely on what participants themselves reveal about their own memory, the measures do not provide a direct measure to clarify the debate. However, the general consensus in modern psychology is to focus on the "what" rather than the "why" and investigate situations of practical relevance in which PEI has a particular influence.

The power of text to change memory

People will sometimes falsely recall text because they made an inference when reading. For instance, early research showed that when recalling sentences that appeared to imply something, participants were more likely to recall a verb that had been implied than the verb used in the original sentence (Brewer, 1977). Roediger and McDermott (1995) introduced a new method for inducing false memories using text, known as the *DRM paradigm.* Using this method,

participants are asked to freely recall items from a list of semantically related words, such as *snow, ice, freeze*. Typically, participants falsely recall the critical lure that was inferred but not actually listed—in this case *cold*—and often with a strong sense of recollection (see Gallo, 2010 for a review). Replications have shown the effect to be robust, and the paradigm has also been used to show differences in memory for special samples, such as participants suffering from Post-traumatic Stress Disorder or with histories of sexual abuse (Bremner, Shobe, & Kihlstrom, 2000; Goodman et al., 2011). Although interesting in terms of general memory distortions, the DRM paradigm does not induce false autobiographical memories that are personally relevant.

Loftus moved away from the original misinformation studies to develop a new *memory implantation method*, for examining whether people could come to remember wholly false autobiographical events, rather than misremember small aspects of events (Loftus & Pickrell, 1995). Participants read narratives about autobiographical childhood events, one of which—getting lost in shopping mall—was false. In this study, 25% of participants recalled some details about the event. Successive studies using this false narrative paradigm have implanted false memories of childhood events—such as planting Slime in a teacher's desk—with a mean false recall rate of 33% (Desjardins & Scoboria, 2007; Garry & Gerrie, 2005).

The power of images to change memory

False photographs have also been used to implant false childhood memories. In one study, participants were shown childhood photographs of

themselves and asked to remember details about the events over a series of three interviews (Wade, Garry, Read, & Lindsay, 2002). One of the photographs depicted a fake event—a hot air balloon ride—in which the participant and a family member had been doctored into. By the end of the study, 50% of participants had come to report either complete or partial false memories of the hot air balloon ride. Similar results have since been found for children using the same method (Strange, Hayne, & Garry, 2007).

Other studies have used digitally edited photographs or videos to induce false memories for recent events as opposed to distant childhood events. For instance, participants shown doctored videos suggesting they completed various actions, have been shown to falsely believe they completed these actions and to develop false memories of doing them (Nash, Wade, & Lindsay, 2009). Indeed, one study found that simply showing participants photographs of a completed action without the participant themselves depicted, was enough to make participants falsely claim they completed these actions (Henkel, 2011a). Thus, false images—even without personalisation—appear to be highly persuasive.

However, a recent study found that participants came to falsely remember performing actions—certain steps involved in building a LEGO vehicle—without any suggestive influence at all (Foster & Garry, 2012). Thus, even without the use of images, performing related actions can lead people to infer that that they completed missing actions, similarly to how reading semantically related words can lead people to misremember a critical lure in the DRM paradigm (e.g., Roediger & McDermott, 1995).

It appears then, that although images may not be necessary—or the only means—to induce false memories, they are still a powerful way of doing so. Indeed, images can also add to the power of other persuasive information. For example, adding a true and relevant photograph to a false narrative has been shown to induce more false memories than the narrative alone (Lindsay, Hagen, Read, Wade, & Garry, 2004). This is still the case when the photograph is not even particularly informative. For instance, people are more likely to believe a false claim when it is presented alongside a "nonprobative" photograph that does not lend any support to the claim or provide any additional useful information (Newman, Garry, Bernstein, Kantner, & Lindsay, 2012).

The power of false evidence to change behaviour

So far, the above research has generally looked at how misleading information or evidence can change peoples' beliefs and memories. However, it is important to also investigate how holding these erroneous beliefs and memories can change peoples' behaviour (Smeets, Merckelbach, Horselenberg, & Jelicic, 2005). After all, acting on a false autobiographical belief or memory can have serious consequences in the real world. As discussed in the first part of this chapter, people might falsely confess to committing a crime when they are innocent, or even falsely accuse someone else. For instance, a person who mistakenly believes they have been sexually abused as a child might be motivated to start criminal proceedings against their supposed abuser.

Studies investigating anxiety have shown that holding a belief can lead to a certain behaviour. For instance, patients with Obsessive Compulsive Disorder

(OCD) are motivated to engage in compulsive checking behaviours because they believe that they would be held responsible for the negative event that would likely occur otherwise (Rachman, 2002). Another example can be seen in children who show avoidance behaviour towards animals that they have been told—and believe to be—dangerous (Field, Argyris, & Knowles, 2001; Field & Lawson, 2003). Thus, beliefs can often be an antecedent to overt behaviours.

Research suggests that fabricated evidence can alter not only peoples' beliefs and memories about an event, but also their attitudes and behavioural intentions too. For instance, in one study, doctored photographs were used to change peoples' perceptions of real past public events (Sacchi, Agnoli, & Loftus, 2007). Participants who viewed a doctored photograph of a protest in Rome—made to look more aggressive than in reality—were more likely to remember serious consequences of the event, described it as being more violent and negative and said they were less likely to participate in future protests than participants who saw the non-doctored version. More recent research has extended these findings to false public events (Frenda, Knowles, Saletanc, & Loftus, 2013). Interestingly, participants were more likely to develop a false memory of a political event—when presented with a fake narrative and photograph—if the event was congruent with their pre-existing political attitudes than when it was dissonant.

Although these studies show the link between false evidence and changes in attitudes and behavioural intentions, they do not show direct behavioural consequences for an autobiographical event. The link between false evidence, autobiographical belief and behaviour has only recently been addressed using a

novel *food questionnaire technique* (Geraerts et al., 2008; Laney, Fowler, Nelson, Bernstein, & Loftus, 2008; Scoboria, Mazzoni, & Jarry, 2008). Using this method, participants complete a questionnaire and receive false feedback that they hated or got sick after eating a specific food as a child. Participants have been shown to report false beliefs and memories of being sick or hating a food during childhood, and have subsequently eaten less of the target food than their control-group counterparts when given the opportunity to do so. These findings have since been extended to alcohol as well as different foods, which could potentially have beneficial implications for reducing alcohol consumption (Clifasefi, Bernstein, Mantonakis, & Loftus, 2013). Indeed, one study importantly demonstrated the potential for positive behavioural consequences from using the technique (Laney, Morris, Bernstein, Wakefield, & Loftus, 2008). After being led to believe that they loved to eat asparagus as a child, participants reported increased liking, desire to eat, and willingness to pay more for asparagus.

The behavioural consequences of believing misleading information are most clearly seen in false confession experiments. Kassin and Kiechel (1996) pioneered the *computer crash paradigm* in which participants complete a mundane typing task alongside a confederate. Part-way through the task, the computer is rigged to crash and participants are falsely accused of pressing the ALT key and causing the crash, and thus also causing the loss of data. For some participants, the confederate acts as a witness and provides additional false incriminating testimony. Participants are ultimately asked to sign a confession statement, and another confederate probes them in the waiting room for information. The conversations are later coded to see if participants really

believed they had pressed the key or were merely complying with the experimenter. In Kassin and Kiechel's original experiment, the majority of participants confessed (69%), and a large proportion indicated that they believed they had committed the act (28%). Some participants even proceeded to make up details about how it happened (9%). Importantly, providing false evidence from the "witness" increased participants' willingness to sign the confession statement. Thus the study clearly demonstrates how false incriminating evidence can increase false confessions and lead people to internalise the act.

The study has since been replicated with the experimenter also acting as a witness and achieved similarly high confession rates (Horselenberg, Merckelbach, & Josephs, 2003). High rates were also found in replications with both adults and children, even when the consequences of confessing were elevated, for instance by being told they would need to return to re-enter the lost data for up to 10 hours (Horselenberg et al., 2006; Redlich & Goodman, 2003). Even simply referring to false and potentially incriminating evidence—a technique known as the "bluff"—can produce comparable confession rates in this paradigm to actually presenting the false eyewitness evidence itself. In one such computer crash study, the experimenter "bluffed" by claiming that another computer held a record of the keystrokes made by the participant before the computer crashed (Perillo & Kassin, 2011). Self-report data showed that participants were more likely to confess, because—somewhat paradoxically—they viewed the bluffed evidence as a means to future exoneration.

The results of these computer crash studies have been extended to show that the presentation of fabricated doctored evidence can also have behavioural

consequences. For example, Nash and Wade (2009) accused participants of cheating on a gambling task and either told participants that a video of the cheating existed or showed them a doctored video of the event. Participants were asked to sign a confession statement and similarly to Kassin and Kiechel's (1996) method, a confederate in the waiting room then probed them to see if participants really believed they had cheated, and thus came to internalise the act. Overall false confession rates were high at almost 100%, and participants who saw the doctored video were more likely to falsely confess to the cheating, internalise the act, and make up details about how they cheated than participants who were simply told the video existed. Thus, the study shows that even an inference to fabricated digital evidence can induce false confessions and beliefs, and that the effect is exacerbated when the evidence is actually presented to the participant. The results of this study have since been extended to show that fabricated video evidence can also induce participants to falsely accuse a confederate of cheating (Wade, Green, & Nash, 2010).

Making a false claim about technical evidence has been found to result in higher false confession rates than making a false claim about eyewitness testimony. For instance, in one study, students accused of exam fraud were more likely to falsely confess when they were told there was an incriminating video than when they were told there was a witness (van Bergen, Jelicic, & Merckelbach, 2008). Interestingly, there was also a correlation between memory distrust and false confessions, such that participants who questioned their memory were more likely to confess. More recent research complements these findings to show that people who distrust their memory are more likely to accept

misinformation than people who hold optimistic beliefs about their memory (van Bergen, Horselenberg, Merckelbach, Jelicic, & Beckers, 2010).

The plausibility debate

People sometimes claim to remember highly unlikely events, such as being abducted by aliens (Clancy, McNally, Schacter, Lenzenweger, & Pitman, 2002). Thus, it is not surprising that studies have shown that it is possible to implant entirely false autobiographical memories of highly implausible events, such as witnessing demonic possession as a child (e.g., Mazzoni, Loftus, & Kirsch, 2001). For instance, several studies used a simple flyer advertisement to implant the impossible childhood memory of meeting Bugs Bunny—a Warner Bros character—at Disneyland (e.g., Braun, Ellis, & Loftus, 2002; Braun-LaTour, LaTour, Pickrell, & Loftus, 2004). Another study found that in children, it was just as easy to induce a false memory of an implausible event of being abducted by a UFO as a plausible event memory of almost choking on sweets (Otgaar, Candel, Merckelbach, & Wade, 2009). Other studies have shown that people can have memories for events that they do not actually believe happened, and indeed know to be highly implausible or even impossible, such as a memory of seeing Santa Claus as a child (e.g., Clark, Nash, Fincham, & Mazzoni, 2012; Mazzoni, Scoboria, & Harvey, 2010). People genuinely feel as if they are "reliving" the impossible event, because the content of the recollection—emotions, perceptual details, coherence—is the same as for a memory of a real event.

In Kassin and Kiechel's (1996) study discussed above, participants were more likely to falsely confess to the crime when it was made to appear more

plausible. For instance, some participants were instructed to type quickly and some to type slower. Those participants in the fast-pace/witness condition were the most likely to sign the statement. Interestingly, in a follow-up study, participants were found to confess irrespective of plausibility but were only likely to internalise the act when it was plausible (Horselenberg et al., 2006). Thus, these confession studies show that plausibility appears to sometimes be linked to both compliance and belief. Indeed, giving participants information that increases the perceived general plausibility of an event—even an implausible event—increases ratings of likelihood that the event personally happened them (Mazzoni et al., 2001). This effect is boosted by providing participants with feedback suggesting that the event has probably happened to them, which increases perceived personal plausibility. Thus, even a small increase in perceived plausibility of an event facilitates a misleading suggestion. However, increasing general plausibility only appears to induce false autobiographical beliefs for relatively benign childhood events—such as having a bone density scan—rather than for an event with personal implications—such as having a rectal enema (Scoboria, Mazzoni, Kirsch, & Jimenez, 2006).

The role of plausibility therefore appears to be relatively complex. If the false evidence is strong enough, people might falsely confess to—and develop false autobiographical beliefs and memories for—an implausible or impossible event (e.g., Braun et al., 2002; Braun-LaTour et al., 2004; Kassin & Kiechel, 1996). However, increasing the perceived plausibility of the event makes this more likely.

Chapter Summary

In the real world, manipulative police interrogation techniques can lead people to falsely confess to crimes they did not commit, and sometimes to internalise the act. It is unknown how often people falsely confess, but prevalence estimates are surprisingly high and the general public have a tendency to estimate higher rates than criminal justice professionals. Worryingly, neither group appears to causally link coercive interrogation techniques with false confessions, and thus do not fully perceive the dangers of wrongful conviction. One interrogation technique that is frequently used is the disclosure of evidence, and sometimes false evidence. This tactic provides suspects with key information about the crime that has not been made public, and this information has been found to be particularly influential when juries decide on a verdict.

False confessions, beliefs and memories have been studied in the laboratory using a variety of different methods. Use of leading questions and presentation of false evidence in the form of images, text, or both, have been shown to change peoples' beliefs, memory and sometimes behaviour for both plausible and implausible events. The next chapter looks at why these changes might occur, and how memory theory can help us to understand and predict situations in which a person might falsely confess to committing a crime.

Chapter 2: Theoretical mechanisms

"He who loves practice without theory is like the sailor who boards ship without a rudder and compass and never knows where he may be cast."

—Leonardo da Vinci (1452-1519)

The previous chapter discussed what is known about false confessions, beliefs and memories in both the real world and laboratory contexts. What are the mechanisms behind false confessions though? How does a person come to believe—and sometimes remember—doing something they did not do? Why exactly is false evidence so persuasive? This chapter outlines four of the main theories that can be used to explain the false evidence effect. Throughout this thesis, these theories will be used as a basis for understanding the findings of a series of four experiments investigating different factors that might influence the power of false evidence.

The Source Monitoring Framework (Johnson, Hashtroudi, & Lindsay, 1993; Lindsay, 2008)

Outline of the framework

According to the *Source Monitoring Framework* (*SMF*), rather than being clear records of events, memories consist of mental events—thoughts, feelings, fantasies, dreams—that are evaluated in the context of current goals, pre-existing knowledge, biases and beliefs. The idea is that memory is not like a video camera

that can replay the event in question at any given time. Memory is supposedly fluid and is reconstructed relative to our thoughts and feelings at the time of access.

The SMF is based on Johnson and Raye's (1981) *Reality Monitoring* (*RM*) model, which proposes that people differentiate between memories of actual events and imagined events by weighing up their average quantitative qualities. Criteria are then used to evaluate the extent to which the characteristics of the mental event match the known characteristics of particular sources. For example, memories of real events are generally richer in perceptual details—sounds, smells, textures—than memories of imagined events. In turn, imagined events generally indicate more *cognitive effort*, and so contain more records of elaborating and identifying (Johnson, Foley, Suengas, & Raye, 1988). Thus when making a *source decision*, a perceptually sparse memory with associated cognitive effort is likely to be judged as being imagined rather than being a memory of a true event. Indeed, recent research shows that true memories tend to be more vivid and require less cognitive effort to generate than intentionally false memories (Justice, Morrison, & Conway, 2012).

The RM model describes how people discriminate between an internally generated source (e.g., an imagined event) and an externally generated source (e.g., an event). However, the SMF extends the RM model by broadening the internal-external discrimination. People may also need to make internal-internal discriminations (e.g., between a dream and a thought) or external-external discriminations (e.g., between comments made by two different people). The SMF also extends the RM model by encompassing other characteristics of memories

besides their quantitative characteristics. Mental events may contain semantic, affective and contextual (temporal and spatial) details as well as perceptual and cognitive information. All of these types of information provide indications as to the source of the mental event.

Although source decisions are often made automatically and without awareness, a more systematic approach might be used if the conclusion conflicts with pre-existing knowledge or beliefs. For example, an extremely rich memory of walking on a ceiling might lead to the conclusion that this is a memory for a real event. However, knowledge of the existence of gravity might then prompt re-evaluation of this memory.

How the framework explains false memory development

A fundamental aspect of the SMF is that false memories for events that did not actually happen are created through the same process as true memories. Because memories are an attribution of a mental event to a particular source, false memories occur when a mental event is misattributed to a real source rather than an imagined one. Source decisions are ultimately reliant on the quality of the information encoded at the time of the event and at retrieval. Thus, any factor that reduces the capacity to encode or retrieve will increase the likelihood of a *source monitoring error,* and a misattribution of a false event to memory. Examples of such factors include strong emotion and divided attention (e.g., Jacoby, Woloshyn, & Kelley, 1989; Mather et al., 2006; Troyer, Winocur, Craik, & Moscovitch, 1999).

The SMF has been used to explain a wide variety of false memory findings. For instance, it is likely that DRM studies—as described in the previous chapter—induce false memories because the associated network that links the words in the list is activated, but the source of this activation is not monitored (Roediger & McDermott, 1995).

The Metacognitive Model (Mazzoni & Kirsch, 2002)

Outline of the model

Mazzoni and Kirsch's *Metacognitive Model* is an extension of the SMF. Importantly however, the Metacognitive Model distinguishes between autobiographical belief and memory. The model also highlights the role of plausibility, such that if an event is judged to be plausible, people are more likely believe that it occurred. Indeed, research suggests that belief, memory and plausibility are nested constructs (Scoboria, Mazzoni, Kirsch, & Relyea, 2004). However, the model recognises that although autobiographical belief and memory are highly correlated, they are still separate entities. Thus it is entirely possible to hold a belief that an event happened, without necessarily having a corresponding memory. This distinction between belief and memory is also made in the *Basic-Systems Model of Episodic Memory* (e.g., Rubin, 2006). In this model, episodic memories are constructed from an interaction between a variety of separate systems (e.g., vision, language, emotion), each with its own properties and neural basis. Importantly, processing in these basic systems is thought to be the basis for different metacognitive judgments, such as belief and recollection.

Mazzoni and Kirch's Metacognitive Model is made up to two parts, depicting how both autobiographical memories and autobiographical beliefs are constructed (Figure 1). The first part of the model outlines how autobiographical memories are created and is a modification of the model of strategic regulation of memory accuracy (e.g., Goldsmith & Koriat, 1999; Koriat and Goldsmith, 1996). In this model, a question initiates a search of our long-term memory for the best *candidate,* which is assessed for its accuracy. If the candidate meets the standard of a pre-determined probability criterion, it will be volunteered as the correct answer. Unlike Koriat and Goldsmith, Mazzoni and Kirsch simply want to know whether an event is attributed to memory, not whether or not people are likely to volunteer or withhold information. Koriat and Goldsmith assume that the search for a candidate will yield only accurate or inaccurate memories or wild guesses. Whereas, like the SMF, the Metacognitive Model assumes that the search could find any type of mental event (thoughts, confabulations, imaginings) rather than already defined memories. These mental events are then evaluated according to pre-existing beliefs and compared to a criterion, which defines how viable the event has to be in order to be classed as a memory.

The second part of the Metacognitive Model depicts how autobiographical beliefs are created in the absence of memory. If a search in memory yields no results, people will generally conclude that the event did not actually happen. However, people may go on to further evaluate the information if the search triggers one of several metacognitive beliefs. For example, people believe that they are more likely to forget an event if time has lapsed since its occurrence, and that they cannot remember anything from the first few years of life due to infantile

amnesia. People also believe that they are more likely to remember an event, the more distinctive or rare its occurrence. If any one of these beliefs is relevant to the mental event in question, inferential processes are then used to decide if the event happened or not. This means that people infer how likely the event is to have occurred by considering pre-existing knowledge and beliefs as well as newly acquired information. This probability is then assessed against a criterion to make the final belief decision.

The two parts of the model are linked because beliefs can be converted into memories. This happens in two stages. First, the assessment of the likelihood of an event occurring influences the memory criterion. If an event seems highly likely, it is more likely to be accepted as a memory. Second, people rehearse the memory (e.g., by imagining it), which enhances the memory content.

How the model explains false belief and memory development

Like the SMF, the Metacognitive Model accounts for how both true and false memories are created through the same process. The model is primarily based on Hyman and Kleinknecht's (1999) theory of false memory creation. According to this theory, in order to develop a false memory of the event, a person must a) accept the event as being plausible, b) construct an image or narrative of the event, and c) make a source monitoring error such that they misattribute the mental event to memory rather than an imagined event. Mazzoni, Loftus and Kirsch (2001) modified this theory to include autobiographical belief as well as memory. In their theory, to form a false memory, a person must a) accept the event as being plausible to have personally happened to them, b)

develop a belief that the event occurred, and c) misattribute the mental event to memory by making a source monitoring error. It is important to note here the emphasis on plausibility. People will supposedly only engage in a search for mental content if presented with what they perceive to be a relatively plausible scenario, and the more plausible it is, the longer and harder people will search (Mazzoni, 2007).

How the SMF and Metacognitive Model explain the false evidence effect

False evidence increases the likelihood of a source monitoring error and thus development of false memory through three main mechanisms: familiarity, credibility and imagery (Nash, Wade, & Brewer, 2009). Because the mechanism of familiarity is discussed in the following section on the Memory Attribution Approach, it is not covered here.

One mechanism through which false evidence might induce false memory is by increasing perceived credibility. At face value, false evidence might appear highly credible and convincing. When evaluating beliefs and memories, people tend to use credibility as an indicator of truth (French, Garry, & Mori, 2011; Horry, Palmer, Sexton, & Brewer, 2012). Thus unsurprisingly, credibility has been shown to be the main factor in determining whether false evidence is accepted or rejected (Nash, Wade, & Brewer, 2009). If the evidence appears credible, people may lower their criteria for believing that the event occurred and thus attribute less vivid or familiar images to memory. Indeed, studies that provide participants with credible misinformation that an event happened to them can create reasonable rates of false belief (e.g., Mazzoni et al., 2001). If this

misinformation is combined with cognitive rehearsal, these beliefs are converted into memories, resulting in high rates of false memory development (see Mazzoni & Kirsch, 2002).

Studies suggest that credibility is much more likely to influence belief than memory. For instance—as discussed in the previous chapter—people are capable of holding a "nonbelieved" memory of an event (e.g., Clark et al., 2012; Mazzoni et al., 2010). Thus, the perception that the event is no longer credible influences the belief that it occurred but not the "reliving" of the mental content.

Another mechanism through which false evidence may work is by promoting imagery. False evidence might contain highly similar details and characteristics to the original event and thus become easily confused. For instance, a doctored photograph is likely to contain high levels of perceptual detail, which is a characteristic usually associated with a real memory. Thus false evidence may promote source confusion and result in source monitoring errors. Indeed, Henkel and Carbuto (2008) showed that participants made more source monitoring errors when they saw true photographs of a completed action than when they read a matching description of the action. Because both forms of evidence should have the same level of familiarity, this suggests that imagery is further boosting the effect of the photograph. The rich perceptual details might also help participants to imagine the event and imagination has been found to induce false claims, beliefs and memories (e.g., Goff & Roediger, 1998; Mazzoni & Memon, 2003; Nash, Wade, & Lindsay, 2009). Indeed, imagined items high in perceptual details are likely to be classed as real, because people tend to quickly

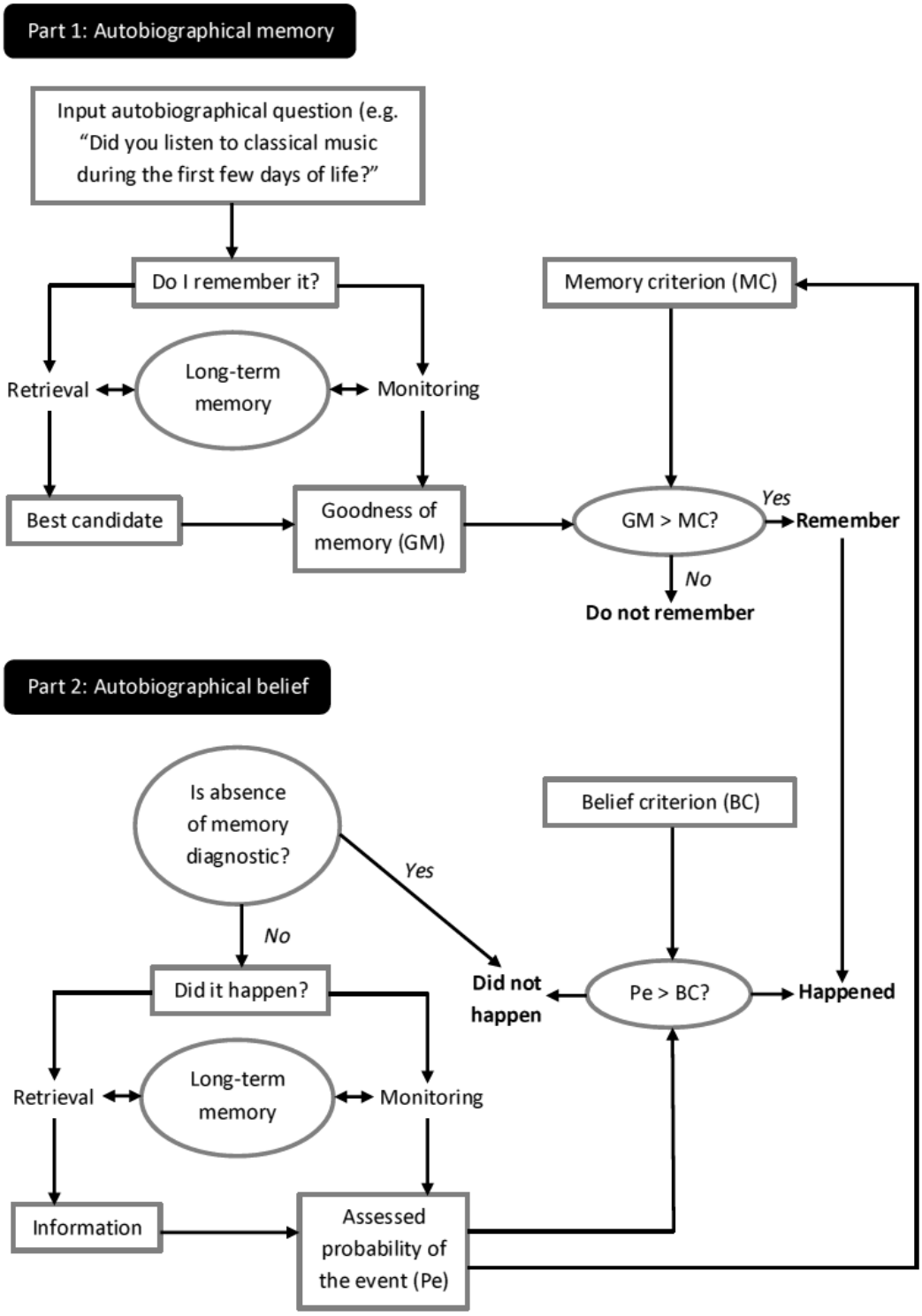

Figure 1. A modified version of Mazzoni and Kirsch's (2002, Fig 6.4, p. 139) diagram depicting autobiographical memory and belief.

base their decision on the idea that "if an item is vivid it is likely to real" (A. Kelly, Carroll, & Mazzoni, 2002).

The Memory Attribution Approach (e.g., Jacoby, Kelley, & Dywan, 1989; Whittlesea, 1993)

Outline of the approach

The *Memory Attribution Approach* is built on the concept of "fluency". When people process a stimulus, they may experience a feeling of fluency, which is a feeling of being easy on the mind to perceive and recall. Like the SMF, people then attribute this feeling to the most likely source at the time, and this is experienced as a feeling of familiarity. If an event in the past seems to be the most likely source, people will attribute the feeling of fluency and familiarity to "remembering" the event. This approach is similar to the SMF because familiarity is viewed as an attribution that is influenced by both the past and present. Fluency is seen as a cue to source, such that mental content that appears "fluent" is likely to be judged as characteristic of a true event. However, the approaches also fundamentally differ, because the SMF views familiarity as only one of a variety of cues to source, whereas the Memory Attribution Approach focuses solely on familiarity.

How the approach explains false memory development

The attribution of fluency usually occurs heuristically and automatically without conscious awareness. Usually this is useful because it facilitates future

interactions with objects, people, or events that have been encountered before. However, sometimes people misattribute feelings of fluency to a plausible source in the past, and this can result in the development of a false memory. These misattributions are what Whittlesea (1993) describes as *illusions of familiarity,* and demonstrate how the current context can influence inferences about memory.

How the approach explains the false evidence effect

False evidence may make an event seem highly familiar, and promote a feeling that the event must have happened. The perceptual details contained in a piece of false evidence are often highly similar to those present in the original event, and thus are processed with a feeling of fluency and familiarity. For instance, studies have shown that images can induce powerful *visual fluency* (e.g., Winkielman, Schwarz, Reber, & Fazendeiro, 2003). This feeling of fluency and familiarity is then misattributed to the past and the person develops a false memory of the event based on the false evidence.

The Self-Memory System Model (Conway & Pleydell-Pearce, 2000; Conway, Singer, & Tagini, 2004)

Outline of the model

The *Self-Memory System Model* (*SMS*) emphasises the importance of the self, and how it is closely related to autobiographical memory. Indeed, the self is thought to modulate the construction of memories. In this way, the SMS model is thought to emphasise the role of the "narrative", and is somewhat related to the

Basic Systems Model mentioned above (Rubin, 2006). In the Basic Systems model, the "self" is thought to have the function of maintaining the stable characteristics of an individual. The self is not seen as a discrete entity, rather a collection of "schema" or organised mental structures spread across several inter-related basic systems. Thus in the Basic Systems Model, the self is neurally dissipated across different areas of the brain that then come together. The SMS model however, focuses less on these discrete systems and neural correlates. The premise of the SMS model is that memory forms a "database" of the history of the self that is motivated by current goals. The SMS has three central components (Figure 2) that come together to form specific autobiographical memories: the *working-self*, the *episodic memory system*, and the *long-term self*.

The working-self consists of goal-based control processes that mediate between episodic memories and the long-term self. The episodic memory system consists of mental content (e.g., perceptual and contextual details) for specific events that when activated, give rise to the feeling of "reliving" the event. Episodic memories are formed by the working-self and used in the short-term to keep a record of goal-based activities. Most episodic memories are forgotten: only goal-relevant episodic memories will become integrated with autobiographical knowledge.

The long-term self comprises the *autobiographical knowledge base* and the *conceptual self*, and serves to hold the information that the working-self draws on to activate goal-directed processes. The autobiographical knowledge base consists of *lifetime periods* (e.g., when I was a young child) and *general events* (e.g., I went to a zoo), which integrate with episodic memories (e.g., I saw a

giraffe with a broken leg) to generate specific autobiographical memories. The knowledge base also contains more generalised *life story schema* that place a person's life story in a social context and in doing so aid identity (e.g., I am a modern woman in academia). In turn, the conceptual self comprises abstract general knowledge about the self in a social context that is grounded in more specific episodic memories.

The knowledge base and working-self have a reciprocal relationship, such that the knowledge base constrains and grounds the goals of the working-self so that they are realistic, but the working-self directs access to autobiographical knowledge and thus memory construction, based on goals. For example, older adults tend to show increased recollection for events from their adolescence or early adulthood—termed the "reminiscence bump"—that are consistent with a period of change and identity confusion, a common psychosocial theme at the time (Holmes & Conway, 1999).

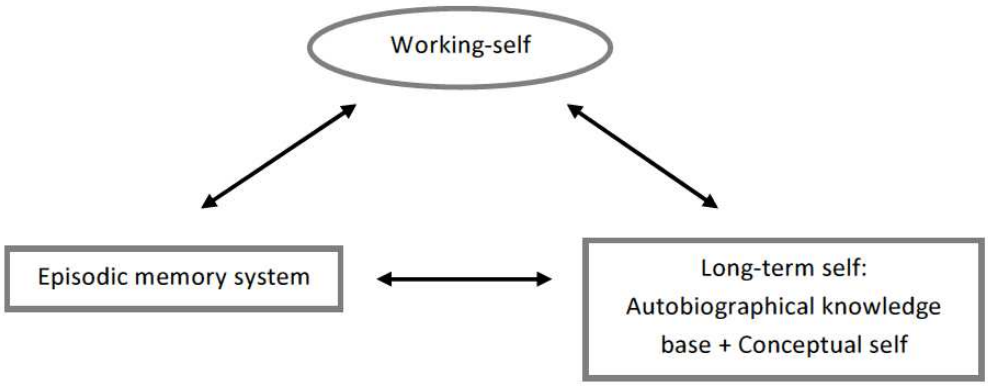

Figure 2. A modified version of Conway, Singer, and Tagini's (2004, Fig 1, p. 494) diagram depicting the Self-Memory System.

Autobiographical memories can either be constructed generatively (voluntary recall) or directly (spontaneous recall) using a combination of the three components. In *generative-retrieval*, a general cue initiates a search of the knowledge base. The working-self directs access to what knowledge is activated and then this knowledge is evaluated using a *retrieval model.* A retrieval model is a control process of the working-self and functions to separate imagined mental content with real memories. Thus retrieval models work to facilitate the source monitoring process described in the SMF. Generative-retrieval is also consistent with the Metacognitive Model, in which a cue triggers a search for mental content, which is then evaluated against a set of criteria. However, the Metacognitive Model is less able to account for involuntary memories than the SMS model. According to the SMS, spontaneous memories occur during direct-retrieval, when a highly specific cue directly activates the knowledge base, without searches being generated and without any initial input from the working-self.

How the model explains false memory development

The SMS explains how false memory can occur through the concept of *coherence*. According to the SMS, any goal change is costly because it uses cognitive and affective resources, has implications for other goals, and means that the self might not function optimally during the transition. Thus the working-self aims to achieve a state of coherence. To achieve this, the working-self may decrease access to memories that challenge current goals, distort or edit memory content, or generate entirely false memories (Conway, 2005). In this way,

experimentally it is possible to induce false memories that are consistent with current active goals. Thus this concept links with work discussed in the previous chapter, showing that people are more likely to develop a false memory for an event that is congruent with their pre-existing attitudes (e.g., Frenda et al., 2013). One of the main ways that a false memory is thought to be generated is through making a source-monitoring error (Jenkinson & Conway, 2012).

The SMS model suggests that people can develop false memories that are consistent with their life story and thus enable coherence. For instance, an adult with an adverse childhood may develop false memories of childhood sexual abuse consistent with their adverse background. Thus the sexual abuse mental content is misattributed from the person's imagination to a series of true events coherent with their adverse life story. Indeed, recent research shows that false memories are often created by recalling and "editing" true memories (Justice et al., 2012). This idea is supported by diary studies, which have shown that people are more likely to accept that a false autobiographical event happened to them if the event contains details from a previous diary entry (Burt, Kemp, & Conway, 2004; Odegard & Lampinen, 2004). It is debatable whether this effect occurs because autobiographical memories consist of separate elements that integrate and "bind" together, or because the previously encountered details promote a familiarity mechanism (e.g., Conway & Pleydell-Pearce, 2000; Jones, Jacoby, & Gellis, 2001).

How the model explains the false evidence effect

The SMS brings together the SMF, the Metacognitive Model and the Memory Attribution Approach to explain how false evidence can lead to the development of false memories. In the SMS model, the working-self works to achieve a state of cognitive consonance, and avoid a state of cognitive dissonance. In this case, cognitive dissonance refers to a feeling of unease that occurs when there is a conflict between autobiographic details and the working-self. False evidence may contain autobiographic self-relevant information (e.g., a picture of a family member, a written reference to a favourite holiday destination) that is interpreted as familiar and "fluent". Because the information is consistent with the working-self—and thus creates a feeling of "ease"—it triggers a broad search of the knowledge base. Because the search is wide as opposed to narrow, it is more likely that any mental content associated with the false evidence will be found and accessed. The false event depicted by this false evidence will then likely be accepted as true and a source monitoring error will occur, in which the false event is misattributed to memory.

However, if the false evidence contains information inconsistent with the working-concept that is interpreted as unfamiliar and "disfluent", a much narrower search of the knowledge base will be made. This then makes it more likely that the evidence will be rejected as false and the event will not be attributed to memory. The narrowest search of all will most likely occur when the false evidence contains both consistent and inconsistent self-relevant information because the working-self will experience a state of conflict and dissonance. Thus false evidence is much more likely to induce a false memory when it contains mostly familiar self-relevant details.

Chapter Summary

The SMF, Metacognitive Model, Memory Attribution Approach and SMS Model can all be used to explain how people come to develop false memories for events and why false evidence can be so powerful. The theories are not necessarily independent of each other, rather they complement each other and can be used to explain different facets of the false evidence effect. The next chapter outlines how these theories bring about research questions regarding the power of the false evidence effect, and examines the need for a new methodology in the field.

Chapter 3: Research outline

"There is nothing impossible to him who will try."

—Alexander the Great (356-323 BC)

The theories discussed in the previous chapter—in particular, the SMF (e.g., Johnson et al., 1993), the Metacognitive Model (Mazzoni & Kirsch, 2002) and the SMS Model (e.g., Conway & Pleydell-Pearce, 2000)—raise a significant theoretical question: does the timing of false evidence mediate its striking effect? For instance, is false evidence more powerful after a very brief delay than when it is presented immediately after the event?

False evidence is powerful because it confuses memory for a source. According to the Source-Monitoring Framework (SMF), people reconstruct past events using the information they have available in the present (Johnson et al., 1993). Specifically, people decide whether their mental products—images, thoughts or feelings—stem from genuine experience or from mental activity such as imagination, fantasy or dreams. To make this decision, people usually evaluate their mental products on multiple dimensions (e.g., familiarity, consistency), automatically and without awareness. But sometimes these decisions go awry and people decide that false autobiographical experiences are real.

Mazzoni and Kirsch's (2002) Metacognitive Model further posits that when people encounter suggestions that contradict their beliefs and memories, including false images, they reconsider the characteristics of their mental products. If their beliefs and memories do not meet the required criteria, people

turn to the information in their environment—which may or may not be accurate—to confirm what happened. The model specifies however, that people will only go on to evaluate and scrutinise their mental content if one or more of several pre-existing beliefs is relevant. Importantly, one of those beliefs is that "memory fades over time". Thus, people would be more likely to turn to external evidence to tell them what happened after a delay, because their memory of the original will have begun to fade. Indeed, most laypeople believe that memory loss for an event is initially rapid and levels out over time—consistent with scientific research (Desmarais & Read, 2011; Magnussen & Melinder, 2012)—and thus that the accuracy of memory for an event decreases the longer ago the event occurred (e.g., Rubin & Siegler, 2004).

The SMS Model (e.g., Conway & Pleydell-Pearce, 2000) also predicts that people will be more likely to believe false evidence after delay. Because the memory for the event will likely have faded, fewer discrepancies are likely to be detected between their original memory and the false evidence. This means that false evidence presented after a delay is less likely to lead to a conflict within the working self than false evidence presented immediately, and thus a broader search for associated mental content will be triggered, and the false evidence is more likely to be accepted as true.

These theoretical premises are supported by research showing that misinformation in general is more powerful when presented after a delay, close to the memory test, rather than immediately following an event (Frost, 2000; Loftus et al., 1978; van Bergen et al., 2010). However, much of the delay effect could rest on the length of the delay used. If the delay is very short, memory of the

original event should likely remain strong and people will not need to turn to external evidence to tell them what happened. They will not need to verify their beliefs about the event by turning to an external source, because they have no reason to doubt their own internal memory. If so, people who see false evidence after such a short delay should be no more likely to turn to—and accept—the false evidence, than those who view the evidence immediately.

The question of the effect of timing of evidence is not only importantly theoretically, but also practically too. Evidence disclosure is a common tactic used by the police—see Chapter 1—and recent research prompts a shift from presenting evidence early on in the interrogation process to late, or gradually throughout, to increase investigators' veracity judgments (Dando & Bull, 2011; Hartwig et al., 2006; Sandham et al., 2011). Importantly however, little is known about the effects of this timing on innocent suspects and their likelihood to falsely confess, believe or remember the crime. This issue is discussed in more depth in Chapter 10.

No research to date has been able to fully test the effect of the timing of presentation of false evidence. This is simply because the current methoologies do not allow for the evidence to be presented immediately, and thus all current false evidence studies involve some form of delay. The current methods can only be used to test differing lengths of time delay rather than the impact of a delay vs. no delay. Often, the methods require testing over multiple sessions. For example, in Nash and Wade's (2009) study, participants needed to attend one morning session and one afternoon session with a 2-3 hour gap in between for the video editing. It is thus impossible to show participants fabricated evidence within a very short

time frame after the event. Multiple sessions also often result in attrition and importantly, can mean that participants' memories of the event become contaminated over sessions.

This thesis introduces a new and simple method for testing the effects of false evidence on belief. Previous methodologies used to study rich false autobiographical belief and memory phenomena—as described in Chapter 1—have yielded impressive results that clearly demonstrate the power of false evidence. However, the methods are generally relatively time-consuming and expensive to conduct. The misinformation method (Loftus et al., 1978), the memory implantation method (Loftus & Pickrell, 1995), the food questionnaire method (Bernstein, Laney, Morris, & Loftus, 2005) and the computer crash method (Kassin & Kiechel, 1996) are all useful procedures, but they often only produce one data-point per participant and require the help of confederates, lengthy protocols and personalised materials that take time to produce.

There is currently no "easy" way of manipulating the effects of fabricated evidence within one simple paradigm. This makes any potential comparison across studies hard, and it is impossible to tell why there is variation in false confession, belief and memory rates. For example, false memory rates in food questionnaire studies tend to be quite low (e.g., only 3% of participants developed a memory of getting sick as a child after eating egg salad in Geraerts et al., 2008). Yet in memory implantation studies some false memory rates are surprisingly high (e.g., up to 50% of participants developed a memory of a childhood hot air balloon ride in Wade et al., 2002). Is this because of the type of memory being induced? Is it because of the methodology? Is it because of other techniques being

used in the studies, such as guided imagery over several interviews in Wade et al. (2002)? Because of these differences and no simple and controlled way of varying factors across studies, very little is known about how the false evidence effect might be moderated.

The new *driving task* procedure was developed to address some of these problems and provide a complementary method to further investigate the false evidence effect. The procedure can be conducted in a single session, and does not require confederates or the construction of personalised materials. The driving task procedure can also easily be adapted to explore other false evidence effects, which will be demonstrated in succeeding chapters. The new method was initially used in Experiment 1 of this thesis to address the opening research question regarding the timing of evidence, and whether any delay effects can be seen when the delay used is strikingly short. The results from Experiment 1 informed theory and raised other questions about the false evidence effect, which were then addressed in later experiments using the new method. The following section outlines these experiments and their aims.

Outline of experiments

This PhD programme uses fabricated evidence to induce both coerced-compliant and coerced-internalised false confessions. The overarching aim is to investigate a series of factors that might mediate or moderate the influential effects of false evidence on autobiographical belief and memory. The thesis is made up of four experiments, each testing a different factor, and all of which use a

novel experimental driving task methodology. This new method is described and evaluated in the next chapter and presented with initial pilot data.

Experiment 1 starts by addressing the question of timing of false evidence using the new methodology. Is false evidence presented immediately after an event more or less powerful than evidence presented after a delay? This research question is extended in Experiment 2, whereby the effects of repeating false evidence—and repeating it over time—are tested. Experiment 3 looks at what happens when different types of false evidence are presented. Is a false photograph more or less persuasive than a piece of text? These different modalities are then investigated further in Experiment 4. Does it matter which piece of evidence is presented first? Are there order and anchoring effects of false evidence? Chapter 8 then discusses the findings of a questionnaire that looks at general attitudes towards digital alteration. Finally, the experimental findings are used to suggest modifications to current memory theory in Chapter 9, and the broader practical, methodological, and theoretical implications are considered in the General Discussion in Chapter 10.

Chapter Summary

This thesis uses a novel method to examine different factors influencing the false evidence effect, and in doing so brings together—and suggests modifications to—current theoretical understanding. The findings from this thesis have potential theoretical, practical and methodological implications. The next chapter outlines the new experimental procedure and how it was used in Experiment 1.

Chapter 4: Experiment 1

"I must govern the clock, not be governed by it."

—Golda Meir (1898-1978)

The previous chapter highlighted the need for a new methodology to address current theoretical and practical questions regarding our understanding of the false evidence effect. This chapter now outlines the novel method implemented throughout this thesis and how it was used in Experiment 1. The primary aim of Experiment 1 was to explore whether doctored evidence is more powerful after a brief delay than when it is presented immediately after the event. Other current methods would not allow for false evidence to presented immediately after the event, and thus could not fully investigate the effects of immediate vs. delayed false evidence. The new computer software was designed to automatically and instantaneously generate the false evidence, and thus eliminated laborious doctoring of materials by the experimenter and the necessary delay required.

In this experiment, participants were wrongly accused of cheating on a driving task and were shown a compelling false video of the cheating either immediately or after a delay. This chapter outlines the finalised methodology used in Experiment 1 in full, and discusses the possible theoretical, practical and methodological implications of the results.

EXPERIMENT 1

Method

Participants

Six progressive versions of the computer software were initially piloted on 37 participants in total, aged 18-23 years (*M*= 18.84, *SD*= 1.17), who took part for course credit. Participants were randomly allocated to the *control* (*n* = 10, *M* = 19.02 years, *SD* = 1.55), *early-evidence* (*n* = 13, *M* = 18.69 years, *SD* = 1.18), and *late-evidence* conditions (*n* = 14, *M* = 18.71 years, *SD* = .83).

The finalised version of the software was then used with a new subset of participants. Seventy-five 18-50 years-olds from Warwick University were recruited to take part in a driving task for credit or £3. Participants were randomly allocated to the control (*n* = 25, *M* = 21.84 years, *SD* = 6.76), early-evidence (*n* = 25, *M* = 20.84 years, *SD* = 2.76) or late-evidence conditions (*n* = 25, *M* = 19.80 years, *SD* = 2.58).[2]

Finalised Materials and Procedure

Participants individually completed a 20-minute hazard-perception test. Onscreen instructions informed participants that they would view 14 video clips, plus a practice clip, of driving situations. These video clips ranged from 15-95 s in length (*M* = 46.27 s, *SD* = 20.40). Participants could score points by clicking the mouse when they saw *hazards*, defined as, "something which causes the driver to slow down, stop or change direction." When participants clicked, a flag appeared at the bottom of the screen. Participants were told to only click when the traffic

[2] An additional 31 participants who cheated on the bonus clip, and four who expressed suspicion were excluded from analyses.

light in the top right corner was green (Figure 3). Clicking when the light was red would be classed as cheating, which participants were told would be taken seriously and may result in disqualification from the experiment. Participants were told that they would be notified during the test when a "bonus" clip was about to appear, and if they identified all the hazards in that clip and scored the highest out of all participants, they would win £50. This monetary incentive encouraged participants to pay attention to the crucial bonus clip, which was always Clip 3.

Figure 3. Screenshot of hazard perception driving test with associated traffic light.

Immediately after the bonus clip, all participants were falsely told onscreen that they had been disqualified because they had clicked during a red light, and thus would be disqualified from the experiment. Control participants were not shown any false evidence. Early-evidence participants were shown false

evidence—a supposed "replay" of their clicking. The clip was replayed with red flags representing the participant's clicks. For a minimum of one click and a maximum of five clicks, the software overlaid a red light where there had been a green light, making it appear as if the participant had clicked during a red light (Figure 4). Early-evidence participants then continued with the driving test. Late-evidence participants continued for a further 9 mins, and were shown the false evidence at the end of the test. There were no significant differences across conditions for the total number of clicks made during the bonus clip, $F(2, 72) = 0.23$, $p = 0.80$, or for the number of clicks overlaid with a red light, $F(2, 72) = 0.88$, $p = 0.42$.

Finally participants were interviewed—supposedly to provide feedback about the test—by the experimenter who was blind to participants' conditions. The experimenter read 10 statements aloud and participants indicated their agreement using a 5-point scale from 1(strongly disagree) to 5(strongly agree). Two statements asked participants how useful they found the test ("Hazard-perception tests like this are helpful to learner drivers") and eight statements asked about their performance on the test ("I performed consistently across clips"). The critical statement "I believe I cheated on the bonus clip", served as a self-report measure of false belief. Participants also described or explained what happened as an additional measure of belief and justification. Participants were probed for details about what happened in the clip if they could not initially remember which clip they were disqualified for. The statement was rephrased to "I believe I clicked on a red light in the bonus clip/clip 3" when necessary. Participants were debriefed and asked to provide retrospective consent.

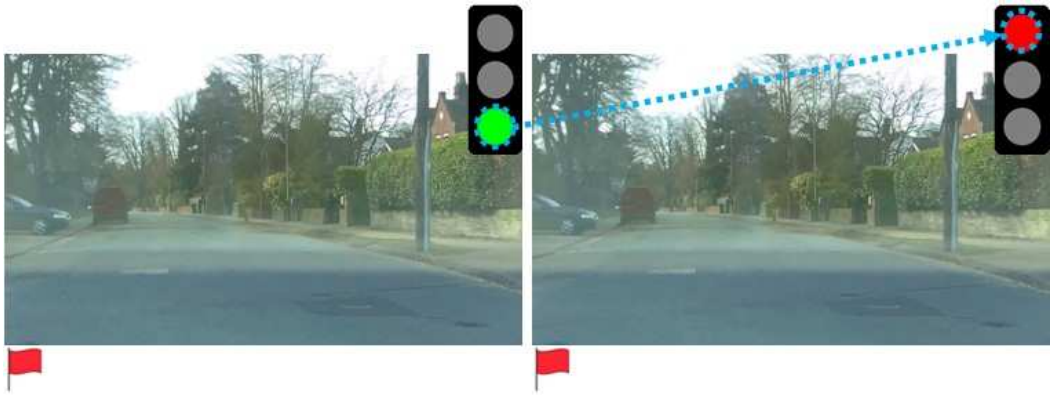

Figure 4. Process of doctoring the replay.

Changes to the procedure during piloting

Although the basic framework did not change, the procedure was modified in four ways at various stages throughout the piloting process. First, the cheating criterion was changed three times. In the first four versions of the software, participants were told that clicking "repeatedly and indiscriminately" would be classed as cheating. However, on average participants only clicked five times during the bonus clip, and thus were reluctant to admit they had clicked "repeatedly in and indiscriminately". In the fifth version, cheating was classed as clicking "too many times". To encourage a sufficient amount of clicking on the bonus clip, the instructions were modified to indicate that the number of clicks allowed was relative to the number of hazards present. However, the modified criterion proved subjective and arbitrary. Participants did not construe their clicking behaviour to involve "clicking too many times". The final cheat criterion used involved implementing a traffic light icon in the top right corner of the screen. Participants were told that this would help them to identify hazards. They could only click on a hazard when the light was green, and clicking during a red

light would be classed as cheating. The superior criterion allowed for an objective accusation, because clicking on a red light was a clear violation of the rules.

Second, the critical question was rephrased according to the cheating criterion modifications. Participants were initially asked to rate their agreement on a scale from 1(strongly disagree) to 5(strongly agree) with the statement "I clicked too many times on at least one clip." To avoid confusion in cases whereby participants thought they might have cheated on a clip other than the bonus, the statement was rephrased in version four to "I clicked too many times on the bonus clip." However, this statement was ambiguous as to whether participants were merely complying with the suggestion that they had cheated, or whether they actually believed they had cheated. The statement was amended in version five to "I believe I cheated on the bonus clip", which specifically measured the extent to which participants falsely believed they had cheated.

Third, information given to participants about the bonus clip varied. In the first five versions of the software, data from both *true-cheaters*—those who actually did cheat on the bonus clip—as well as *false-cheaters*—those who did not cheat but were wrongly accused—were collected. Thus, the experiment had a 2 x 3 design, with cheater type (true, false) and timing of evidence (control, early-evidence, late-evidence) as independent variables. In the true-cheater condition, participants were encouraged to identify all hazards to obtain the highest score, and to click several times per hazard to ensure their click was registered. The instructions in version four added that the clip would only be positively marked, such that any clicks participants made on non-hazards would not be counted against them. The aim was to induce participants to actually cheat in the task, to

obtain a sample of true-cheaters. In the false-cheater condition, participants were encouraged to be as accurate as possible by clicking the fewest number of times necessary to identify the hazards. The instructions in version four added that the clip would be negatively marked, such that clicks participants made on a non-hazard would harm their score. The aim was to discourage participants to cheat so that they could then be falsely accused of cheating later on. However, the incentive to cheat in the true-cheater condition was not strong enough to induce participants to break the rules. Out of the 17 participants in the true cheat condition, only four actually cheated on the bonus clip. Thus, a large sample of participants had to be excluded because they had not actually cheated. I therefore decided to only have a false cheater condition, which followed the original aim to investigate the effects of moderating factors of false evidence on innocent people. In version six of the software, all participants were given a selfish incentive. They were told that they could win £50 if they scored the highest out of all participants on the bonus clip. This served to ensure that participants paid attention to, and remembered, the bonus clip.

Fourth, the feedback session was modified. During piloting, all participants completed a written feedback form. However, the responses provided using this method were inadequate. Participants often failed to read instructions correctly and thus provided unusable data, or missed sections out completely. Thus, a structured interview was implemented in the final method. The questions were identical but read aloud to participants by the experimenter. Participants provided vocal responses and were probed when responses were unclear.

In sum, by the seventh finalised version, all participants were told that they could win £50 by scoring the highest out of all of the participants in the bonus clip. They were then falsely accused of cheating on the bonus clip by clicking during a red light, and some participants were shown false evidence of the supposed cheating. Finally, participants were interviewed and asked to rate the extent to which they believed they had cheated on the bonus clip.

Results and Discussion

The initial pilot sample was too small to conduct meaningful analyses. However, descriptive statistics showed an early trend. Of the 20 participants who were falsely accused of cheating, late-evidence participants clearly showed the highest rates of belief (M=3.38, SD= 1.06), compared to early-evidence (M= 2.86, SD= 1.86) and control participants (M= 2.60, SD= 1.52) participants. The finalised method was then used on the new participant subset and the results are detailed below.

Importantly, did delaying the false video influence participants' beliefs and justifications? First, participants' self-reported belief ratings were analysed (Figure 5). A Kruskal-Wallis test showed a significant effect of condition on belief ratings, $\chi^2(2, N = 75) = 12.29$, $p = .002$, $\eta^2 = .17$.[3] Follow-up Mann-Whitney U tests revealed that late-evidence participants reported higher belief scores than early-evidence participants (average $rank_{late} = 31.12$, average $rank_{early} = 19.88$; $z = -2.86$, $p = .004$). Late-evidence participants also reported higher belief scores than control participants (average $rank_{late} = 31.74$, average $rank_{control} = 19.26$; $z = -3.15$,

[3] Non-parametric tests are more appropriate than a Oneway ANOVA with posthoc tests because the ordinal data violate normal distribution and homogeneity of variance assumptions.

p = .002). Belief ratings did not differ significantly between early-evidence and control participants, *p* = .409.

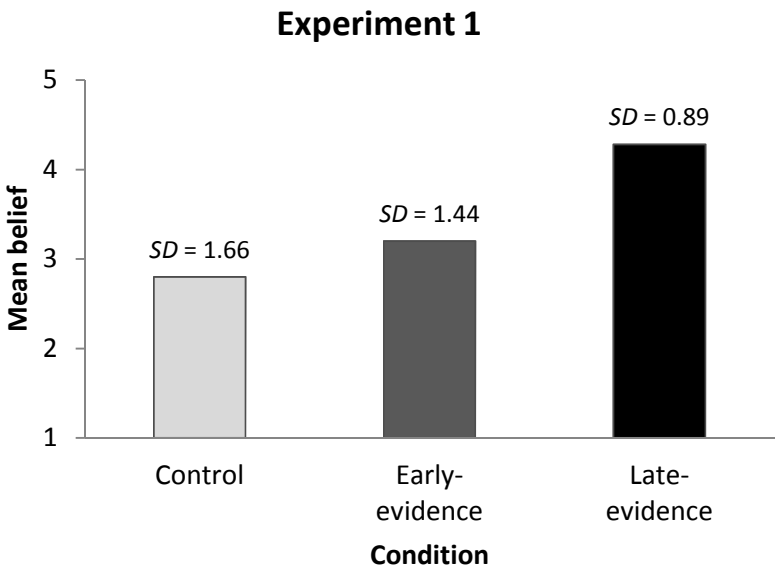

Figure 5. Mean self-reported belief ratings as a function of condition in Experiment 1.

Two judges also used Redlich and Goodman's (2003) criteria to determine which participants believed they cheated. Redlich and Goodman (2003) used the computer crash paradigm described in Chapter 1 with participants from different age groups. They assessed whether or not participants internalised the act by coding participant responses for evidence of *No, Partial* or *Total internalisation*. The same criteria were used in this thesis to determine if participants actually believed that they had cheated, which is termed "belief" rather than "internalisation" throughout. The judges were thoroughly trained and blind to participants' conditions. They classified participants' explanations for the cheating as reporting *No, Partial* or *Full belief*, and agreed on 89.3% (k¼.83) of

categorizations. Participants were classified as partially believing if they reported that they might have cheated, such as, "I clicked on a red accidentally I guess…" Participants were classified as fully believing if they reported that they definitely did cheat, such as, "I clicked when there was a red light…" Disputed cases were placed in the more conservative category.

Table 1. *Judges' belief ratings by condition in Experiment 1.*

	Judges' belief ratings		
	No	Partial	Full
Control participants	36%	32%	32%
Early-evidence participants	20%	48%	32%
Late-evidence participants	12%	20%	68%

Table 1 shows judges' categorizations of participants' beliefs. Full beliefs were significantly more common in the late-evidence condition than in the early-evidence, $\chi^2(1, N = 50) = 6.48$, $p = .011$, Cramer's V = .36, and control conditions, $\chi^2(1, N = 50) = 6.48$, $p = .011$, Cramer's V = .36, but did not differ significantly between early-evidence and control participants, $p = .100$. When collapsed across partial and full belief categories, judges rated significantly more late-evidence participants as believing than control participants, $\chi^2(1, N = 50) = 3.95$, $p = .047$, Cramer's V = .28. Early-evidence participants did not believe significantly differently to late-evidence or control participants, both $ps \geq .208$. This is likely due to the high levels of partial belief displayed by participants in the early-evidence condition, which appears to reflect their uncertainty in the accusation.

Indeed, early-evidence participants did not significantly believe more than control subjects on either the self-report or judges full or collapsed belief ratings. Thus both self-reported and judges belief measures showed that a delay is necessary before false evidence is presented for participants to believe the false accusation.

Judges also used a modified version of Redlich and Goodman's (2003) criteria to classify participants as reporting *No*, *Partial* or *Full confabulation.* Participants were judged as partially confabulating if they speculated about what might have happened. This was often—but not always—indicated by hedging within the sentence ("I was looking at the screen. Maybe I didn't notice it"). Participants were judged as fully confabulating if they described how the cheating occurred ("I wasn't looking at the lights, I just didn't realise.") Judges agreed on 81.3% (k¼.69) of categorizations and once again the more conservative category was used in disputed cases.

However, on further inspection, it became apparent that participants were not necessarily confabulating details, rather justifying their actions. Participants tended to make up excuses as to why they committed the act—and how it fitted with their life script—rather than actively embellishing their memory of the event. Although confabulation generally infers the existence of a false memory—and indeed over time has been shown to lead to false memory development (e.g., Drivdahl, Zaragoza, & Learned, 2009; Zaragoza et al., 2001)—this is not necessarily true of justification. The two measures are related but not synonymous. Indeed, confabulation may actually be prompted by an inability to justify a moral judgment (e.g., Cushman, Young, & Hauser, 2006). Thus,

throughout this thesis, coded confabulation has been relabelled to *No, Partial* or *Full justification.*

It is important to note however, that full justification responses also often demonstrated confabulation. For instance, the statement "...I was focussing on the screen and then it changed, I saw it change," clearly indicates that the participant has developed a false memory of the traffic light changing to red, and has then embellished the memory. However, this was not always the case for full justification responses. For instance, the statement, "...Because I was so focussed looking at the hazard that I didn't look at the lights," does not necessarily imply an actual false memory of the cheating, rather just an explanation as to why it happened. Thus, although there was often an overlap between full justifications and confabulation, the term *Full justification* is used throughout this thesis to be conservative.

Table 2. *Judges' justification ratings by condition in Experiment 1.*

	Judges' justification ratings		
	No	Partial	Full
Control participants	72%	0%	28%
Early-evidence participants	60%	20%	20%
Late-evidence participants	40%	24%	36%

In Experiment 1, for full justifications, there were no significant differences between groups, all *p*s ≥ .208 (Table 2). However, when collapsed across partial and full justification categories, judges rated that significantly more

late-evidence participants justified than control participants, $\chi^2(1, N = 50) = 5.20$, $p = .023$, Cramer's V = .32. Early-evidence participants did not justify significantly differently to late-evidence or control participants, both $ps \geq .157$. Thus false evidence must be delayed for participants to justify how the cheating occurred. Most participants who justified said that they did not notice the red light because they were focussing on recognising hazards on the main screen (e.g., "I wasn't looking at the lights, I just didn't realise").

Summary of findings

In summary, the results show that delaying false evidence by a mere 9 minutes enhances its effect. In line with Mazzoni and Kirsch's (2002) Metacognitive Model, participants' memories of the bonus clip may have faded enough in this period to prompt them to turn to external evidence—the fabricated video—for verification. Participants then used the video to construct details about how it happened. Surprisingly, there was a null effect of early evidence. Thus, timing not only moderates the false evidence effect by making it more powerful, but it also mediates its effect. Without a delay, false evidence was no more influential than the control. There must be a delay before the false evidence is presented for it to have any effect at all.

A possible counter-explanation for the pattern of results is that the delay between the false video and the interview, rather than the delay between the target event and the false video, increased false beliefs and justification. This possibility is examined in Experiment 2 by including an *immediate-repeated-test* condition. However, as a check for this experiment, a further 25 adults were tested in a

modified version of the early-evidence condition (n = 25, M = 24.24 years, SD = 7.10). Participants encountered the false evidence immediately after the target event, similar to the original early-evidence group, but were interviewed immediately after this; thus eliminating the video-to-interview delay. A direct comparison of data across conditions is not possible because these participants were not randomly assigned to a condition at the start of the experiment. However, the modified early-evidence participants showed a highly similar pattern of results to the original early-evidence participants. Modified early-evidence participants displayed a mean self-reported belief of 2.80, with judges rating participants as fully believing 28% of the time, and fully justifying 24% of time. These results suggest that the delay between the event and the false evidence, not the event and interview, drives the timing effect.

Theoretical implications

Theoretically the results of Experiment 1 are consistent with Mazzoni and Kirsch's (2002) Metacognitive Model and highlight the importance of intervening metacognitive beliefs. Why are participants more likely to believe they cheated when false evidence is delayed? According to the model, the accusation of cheating initiates a search of participants' memories. Inevitably, participants will not remember the false event. Participants then have to decide if this lack of memory is diagnostic of the cheating not actually having occurred. This decision is influenced by several pre-existing metacognitive beliefs acquired earlier in childhood. One such belief is that forgetting is more common after a delay because memories fade over time (Cornoldi, 1998). Thus the absence of a clear

memory for an event that occurred some time ago does not necessarily mean it did not happen. Experiment 1 suggests that late-evidence participants then turn to inferential processes to decide if the event happened, which take into account beliefs, new information, and existing knowledge. Participants then assess the probability that the event occurred. If participants decide that the event is likely to have occurred based on these inferential processes, they develop a false belief that they cheated. However, this is unlikely to occur for early-evidence or control participants because the "memories fade over time" belief does not account for their inability to remember the event, and thus they decide that the lack of memory is diagnostic of the event having not occurred.

How then, might the false belief of having cheated develop into a false memory? Mazzoni and Kirsch (2002) stipulate that beliefs and memories are separate entities and thus having a false belief does not necessarily indicate the presence of a corresponding false memory. A belief is a judgment or conviction about an event, whereas memories are mental events with a recollective quality. We assess mental products using a pre-determined set of criteria to decide if they should be classed as memory or fantasy. In the case of late-evidence participants in Experiment 1, the participants' earlier assessment that the event probably happened is likely to lower these pre-determined criteria, and thus induce false memories. This is exacerbated by the fact that the false video used in Experiment 1 contains qualities that are highly similar to real memories e.g., vividness, plausibility (Johnson et al., 1993).

According to Mazzoni and Kirsch, false memories are more likely to be created when the memory content is enhanced. The false video enhances the

initial accusation of cheating because it 'borrows' features of the original event (Henkel & Coffman, 2004). For instance, the participant views their exact series of clicks made on the bonus clip once again. These borrowed features are important because they make the event seem familiar, and only familiar supporting details will be accepted as fact (Jacoby, Kelley, et al., 1989). The memory content might then be further enhanced by confabulation. Confabulating details about how the event occurred promotes rehearsal, and thus the accusation repeatedly gains fluency, making it feel as if we are remembering (Jacoby, Kelley, et al., 1989). Indeed, studies have shown that forcing participants to confabulate either details about an event or whole events entirely, leads to false memory development over time (Chrobak & Zaragoza, 2008; Zaragoza, Payment, Ackil, Drivdahl, & Beck, 2001). In fact, combining misinformation—in this case, the false accusation of cheating—with cognitive rehearsal can produce high rates of false memory (e.g., Hyman & Billings, 1998; Wade et al., 2002).

Importantly, the results of Experiment 1 emphasise the role of internal metacognitive factors in developing false beliefs over and above other external factors. Previous studies have shown that the effects of false evidence are moderated by how credible or convincing the evidence is perceived to be (e.g., Nash, Wade, & Brewer, 2009). Indeed, even the perceived credibility of the source of that evidence affects development of false beliefs (Scoboria, Wysman, & Otgaar, 2012). In line with Mazzoni and Kirsch, a highly credible source can cause people to lower the memory criterion they use to decide if something should be attributed to memory or not. Essentially, credibility of a source is used as a proxy for truth judgments (Echterhoff, Hirst, & Hussy, 2005; French et al., 2011;

Horry et al., 2012). However, in Experiment 1 the false video remained the same throughout, and thus its perceived credibility should have remained consistent. Indeed, it is likely that the credibility of the false video was perceived to be consistently high, because it was presented by a trusted authoritative source: a computer. Thus, Experiment 1 suggests that perceived credibility has little impact when false evidence is presented immediately after an event, because early-evidence participants did not simply accept the highly credible evidence as fact. I return to this idea in Chapter 9.

Practical implications

On a practical level, the findings from Experiment 1 raise questions about the disclosure of evidence during police interrogations. It used to be common practice to present evidence to suspects at the start of the interview in an attempt ease them into a confession. By 1996, this still appeared to be the case in the US (Leo, 1996), but not in the UK (Moston & Engelberg, 1993). There has since been a steady shift towards disclosing evidence towards the end, rather than at the start of an interrogation, because studies have shown that this method enhances the interviewer or observer's ability to detect deception (Hartwig, Granhag, Strömwall, & Kronkvist, 2006; Hartwig, Granhag, Strömwall, & Vrij, 2005). Also, as discussed in previous chapters, some US jurisdictions allow the presentation of fake evidence during interrogations, which various training manuals actively encourage (e.g., Inbau et al., 2005). Because of this, using fake evidence may be viewed as an acceptable interrogation tactic, especially if the delay between the target event and disclosure of the evidence is minimal.

Experiment 1 suggests, however, that a mere 9 minute delay can increase the power of fake evidence to elicit false beliefs and justification, and thus increase the risk of false positives.

Recent research has shown that disclosing evidence incrementally (called *phased* or *tactical disclosure*), may be optimal in terms of both increasing veracity judgments and decreasing investigators' cognitive load (Dando & Bull, 2011; Sandham et al., 2011). This method relies heavily on the existence of sufficient amounts of evidence but can be extremely effective. Truthful suspects are given the opportunity to convey their honesty early on and account for each piece of evidence. However, deceptive participants are unable to adhere to a previously created *lie script* when they are continually presented with challenges to their account (Dando & Bull, 2011). The results of Experiment 1 could be extended to investigate how false belief rates are affected by the phased disclosure of several pieces of false evidence.

Methodological implications

Overall, Experiment 1 demonstrates the effectiveness of the new driving task methodology for four reasons. First, participants were convinced by the set-up, with only four participants indicating suspicion. The majority of participants appeared surprised during debriefing and expressed interest in the outcomes of the study ("I didn't think I cheated but then I saw the evidence and thought I must have!"; "I completely believed it!"; "Wow! You were really good! I couldn't suspect anything! This was a great experience!") Second, the two belief measures demonstrate criterion validity because there was a strong convergence between

participants' self-reported ratings and judges' categorizations. Third, the methodology yielded consistent results despite the continual modifications to the six piloted versions. The pattern of findings in Experiment 1 matched the initial pilot findings. Fourth, the results are comparable to other false confession studies. For instance, the 68% full belief rating of late-evidence participants is comparable to the 67% reported by Nash and Wade (2009) for participants who were shown a doctored video of themselves cheating on a gambling task. In the current study, however, the doctored videos did not depict the participants, only the act they supposedly committed. This finding fits with recent research showing that exposing people to photographs of completed actions—photos that do not depict the participants themselves—leads to erroneous claims that they performed these actions (Henkel, 2011a).

Importantly, the false event was designed to be relatively plausible. As Hyman and Kleinknecht (1999) state in their theory of false memory creation, before a participant can come to falsely believe and potentially develop a false memory of an event, they must first view the event as being plausible. The plausibility debate was discussed in Chapter 1, with studies suggesting that although people can come to believe in implausible events, it is more likely when the event is made to seem more plausible. The false cheating accusation in the new methodology was plausible for two reasons. First, as Johnson (2006) suggests, the false evidence was based on real content from the event itself. In the final version, participants were shown a video depicting the exact series of clicks they had made during the bonus clip. Thus the false evidence was actually partially true, such that the series of clicks was identical and only the traffic light

sequence was altered. Second, the driving method essentially involves a dual-attention task in which participants must attend to both the traffic light and the main video screen at the same time. This makes the task much harder, and thus the accusation of cheating appears more plausible because participants could have made a mistake by clicking on a red light whilst attending only to the main screen. Divided attention also means that the event itself is encoded poorly at the time and thus when accessed, the memory for the event is vague, with little information available to help identify the source. According to Mazzoni and Kirsch's (2002) Metacognitive Model, a 'vague' memory that lacks source-defining characteristics is unlikely to meet the memory criterion, and thus participants would likely turn to external evidence to tell them what happened. Previous studies support this premise by showing that divided attention tasks induce recognition errors between old and new items (e.g., Jacoby, Woloshyn, et al., 1989). Dividing attention—and indeed, depleting cognitive resources in general—has been shown to increase adults' false memories of critical lures in the DRM paradigm (Otgaar, Alberts, & Cuppens, 2012; Otgaar, Peters, & Howe, 2012).

It is also important to note here that owing to the plausibility of the event, it is probable that some participants accidentally clicked when they were not supposed to. Thus, when using the new driving task method, there is no differentiation between participants who deliberately click and those who mistakenly click on a red light. The term "cheater" is used throughout this thesis to refer to participants who both intentionally and unintentionally break the rules. For the purposes of this thesis, this differentiation is of minimal importance. Theoretically, participants admitted to—and sometimes believed—that they

clicked on a red light, which in itself is an act of cheating. Practically, in the real world, suspects may also falsely confess to an action that then constitutes a crime. Indeed, police may initially try to coerce a confession by suggesting that the suspect made a mistake.

Although the false event was highly plausible, it had clear consequences for the participant, such that participants thought they would be disqualified from the experiment. Thus, what was gained in experimental control and plausibility, did not necessarily need to be lost in terms of ecological validity. Anecdotally, many participants immediately confessed to having cheated without prompting and apologised to the experimenter. Several participants expressed particular concern regarding the consequences. This suggests that the false evidence manipulation appeared genuine, and that participants were worried about the consequences of having cheated. Thus participants may have been stressed during interviewing with the experimenter, who essentially was in a position of authority. This is not completely removed from a police interrogation, whereby a confused suspect may appear anxious when questioned by a police officer about a crime they have not committed.

At first glance the large number of participants who cheated and were removed seems problematic. To ensure that the accusation of cheating was plausible, participants were accused of something that they could potentially have committed. Thus, a high portion of participants were expected to actually cheat on the bonus clip. However, after initial piloting, the primary goal was to determine the effect of presenting people with evidence that contradicts their actual experience. To this end, only enough data was collected to allow for statistical

analysis for those participants and thus there was not sufficient data to study true-cheaters. It is possible though that this procedure could be modified by researchers in the future to examine true-cheaters. Indeed, previous research has shown that comparing true and false confessors is useful when trying to capture certain key elements of real-world interrogations and confessions (e.g., Russano et al., 2005). This idea is returned to in Chapter 10.

Interestingly, a study conducted by Lindsay (1990) showed that when participants were warned that information given to them was false and thus to scrutinise the source of their memories, they did not report suggested details presented after a delay but they did when the details were presented immediately after the event. This suggests that the time lag actually helped participants to differentiate between the sources and thus these participants were less susceptible to misinformation. In the new driving procedure, during the interview participants were asked for their own opinion about what they thought happened, and thus they were encouraged to think about the source of information. Despite these instructions, participants in Experiment 1 were still more susceptible to the false evidence after a delay. This is the exact opposite of what Lindsay (1990) found and suggests that being given a warning about information being false might help protect against the enhanced effects of false evidence over time. However, studies on post-suggestion warnings show that they have little impact (e.g., J. E. Kelly & Nace, 1994; Nash, Wade, & Lindsay, 2009) and thus a warning before the false evidence is presented might be more effective. The influence of pre-suggestion warnings on the timing of false evidence presentation is a matter for future research.

Following the results of Experiment 1, Experiment 2 was amended in four ways. First, participants viewed a doctored photograph depicting a click during a red light rather than a video. This would provide greater control because all participants would be accused of clicking once rather than being shown between one and five clicks during a red light. Second, the interview included six questions rather than ten. Four bogus questions were removed to shorten the length of the interview. Third, the critical statement was rephrased to measure compliance as well as false belief. Specifically, participants responded Yes/No to the statement "I cheated on the bonus clip" (compliance) and rated their belief in this item on a 7-point scale from 1 = strongly disbelieve to 7 = strongly believe (belief). The belief scale was expanded from five to seven points to allow for greater sensitivity on ratings. Finally, judges no longer categorised participants' statements for belief. These additional ratings were unnecessary due to their high consistency with the self-reported belief ratings in Experiment 1.

Chapter Summary

Experiment 1 successfully used the new driving task method to investigate what happens when the timing of the presentation of false evidence is varied. The findings show that presenting evidence immediately after the event is no more persuasive than the accusation of cheating by itself. However, showing evidence after a brief delay increases its power. Theoretically, the results fit with the Metacognitive Model and suggest that people are more likely to turn to external evidence when their original memory has faded slightly. The results also have practical implications for evidence disclosure during police interrogation.

Chapter 5: Experiment 2

"Right now I'm having amnesia and déjà vu at the same time."

—Steven Wright (1985-Present)

How consistent are the timing effects of false evidence found in Experiment 1? Does false evidence always have a null effect when presented immediately? Is it always more powerful when presented after a delay? Research suggests that repeating false evidence should make it more compelling. Studies have shown for instance, that when deliberately misleading information is repeated, both adults and children can be lead to misremember details of an event they have witnessed (Ceci, Loftus, Leichtman, & Bruck, 1994; Gobbo, 2000; Mitchell & Zaragoza, 1996; Zaragoza & Mitchell, 1996).

However, studies investigating the effects of repetition often involve repeating the information over time, and thus confound the effects of repetition with the effects of timing. Experiment 2 investigates these two variables using the novel driving task procedure.

Aims of Experiment 2

The aims of Experiment 2 were threefold. The first aim was to investigate the effects of repeatedly presenting false evidence. We know that people are generally more likely to judge information as true when it is repeated—the *illusion of truth* effect (Hasher, Goldstein, & Toppino, 1977; Hawkins & Hoch, 1992; Schwartz, 1982), and this holds even if people are initially told the

information is false (Begg, Anas, & Farinacci, 1992). Research shows that participants tend to over-generalise this truth bias, such that one person repeatedly stating an opinion can lead other people to believe that this opinion is prevalent in a group of people (Weaver, Garcia, Schwarz, & Miller, 2007). Repeatedly showing a photograph of a completed action can even lead people to erroneously claim that they performed this action (Henkel, 2011a). Thus repeatedly showing participants false evidence of an event is likely to increase their belief that the event actually happened.

Repetition may work by promoting a feeling of *fluency* or being "easy on the mind" to perceive and recall (Jacoby & Kelley, 1987). Fluency is sometimes confused with familiarity, which people interpret as an indication of "truth" (Alter & Oppenheimer, 2009). Repetition may also make information appear more credible, and credible images are more easily confused with similar real memories (Nash, Wade, & Brewer, 2009). These mechanisms may be linked, with familiarity acting as an indicator of credibility (Foster, Huthwaite, Yesberg, Garry, & Loftus, 2012). For instance, if something feels fluent, people tend to attribute it to a more intelligent source (Oppenheimer, 2006).

The second aim of Experiment 2 was to investigate how repetition and timing of false evidence might work in combination. We already know from Experiment 1 that false evidence is more powerful following a delay, and it is likely that repetition might increase its influence. Research has sometimes found an effect of repetition in a single session (Foster et al., 2012; Zaragoza & Mitchell, 1996); and sometimes not (Warren & Lane, 1995). A delay between repetitions, or between repetitions and testing, may in fact be necessary to observe

an effect at all. Indeed, repeating misleading information after 1 month leads to memory errors (Gobbo, 2000) and memory errors are observed in single sessions when there is a delay between repetitions and testing (Weaver et al., 2007). It is possible that some form of delay might be needed because immediate repetition could induce suspicion. Participants might then scrutinise the false evidence and their original memory of the event. According to Mazzoni and Kirsch's (2002) Metacognitive Model, this scepticism might lead participants to be more stringent with their memory criterion, and therefore reject the false evidence in favour of their own memory of events. Thus although it is likely repetition with a delay might induce high levels of false confessions and belief, it is possible that immediate repetition may in fact have the opposite effect.

The final aim of Experiment 2 was to add an extra condition to clarify which delay in Experiment 1 was causing the false evidence effect. In Experiment 1 it was unclear whether the delay between the event and the evidence, or between the evidence and the interview, was causing the effect. The modified early-evidence condition indicated that it was the delay between the event and the false evidence. However, this condition was not a controlled and randomly assigned condition of the original experiment and thus was not directly comparable. Experiment 2 also tested how consistent the results of Experiment 1 were using the driving task procedure. If timing is in fact highly influential as Experiment 1 found, Experiment 2 should demonstrate the powerful effects of false evidence repeated after a delay.

Thus there were four conditions in Experiment 2. Participants were presented with false evidence once immediately after the event (*immediate-once*),

twice immediately after the event (*immediate-repeated*), or twice over time (*delay-repeated*). An *immediate-repeated-test* group also viewed the evidence twice immediately like the immediate-repeated group, but finished the experiment straight after viewing the false evidence. Comparing the immediate-repeated-test group to the immediate-repeated and delay-repeated groups enabled differentiation between the influence of timing versus repetition. Because delayed evidence was already found to be persuasive in Experiment 1, there was no *delay-once* condition in which the false evidence was seen once at the end of the test. Thus, Experiment 2 does not follow a fully crossed 2 (Repetition of evidence: Yes vs. No) x 2 (Timing of evidence: Immediate vs. Delayed/Delayed over time) design. It does however, include the immediate-once condition, which acts as a control comparison for repeated evidence.

EXPERIMENT 2

Method

Participants

One hundred and twenty adults, 17-44 years old, were randomly allocated to the immediate-once (n = 30, M = 21.70 years, SD = 4.98), immediate-repeated (n = 30, M = 19.97 years, SD = 1.33), delay-repeated (n = 30, M = 20.47 years, SD = 2.26) or immediate-repeated-test (n = 30, M = 21.33 years, SD = 3.06) conditions.[4]

Materials and Procedure

The procedure followed that of Experiment 1 with three minor amendments. First, participants viewed a doctored photograph depicting a click during a red light rather than a video. Second, the interview was streamlined to include six questions rather than ten. Finally, the critical statement was rephrased to measure both compliance and false belief. Specifically, participants responded Yes/No to the statement "I cheated on the bonus clip" (compliance) and rated their belief in this item on a 7-point scale from 1 = strongly disbelieve to 7 = strongly believe (belief). Including the Yes/No compliance measure provided a direct assessment of the frequency of false confessions.

Immediate-once participants viewed the evidence once, immediately after the accusation. Immediate-repeated participants viewed the evidence twice sequentially and then continued with the test for approximately 9 minutes. Delay-repeated participants viewed the evidence once initially and then again at the end

[4] A further 35 participants who were tested but cheated on the bonus clip, and two who indicated suspicion, were excluded.

of the test. Immediate-repeated-test participants viewed the evidence twice sequentially but were then told that because they cheated they could no longer continue and were tested immediately. The total number of clicks made during the bonus clip did not significantly differ across conditions, $F(3, 62.96) = 1.10$, $p = 0.36$.[5]

Results and Discussion

Did repeating evidence make it more persuasive? Was repetition with a delay more powerful than without a delay? To answer these questions, it was first necessary to investigate if there were any differences between the immediate-repeated and immediate-repeated-test conditions. Preliminary analyses revealed that immediate-repeated and immediate-repeated-test participants did not differ significantly on compliance, $\chi^2(1, N = 60) = .11$, $p = .739$, Cramer's V = .04, belief, $z = -.56$, $p = .576$, or justification measures, both $\chi^2(1, N = 60) = .00$, $p = 1.000$, Cramer's V = .00. This confirms that in Experiment 1 it was the delay between the event and evidence that promoted false beliefs and justification, rather than the delay between the evidence and testing. In the subsequent analyses, these two conditions are collapsed and referred to as the immediate-repeated group.

Did repetition of false evidence influence false confession rates? For the Yes/No compliance measure, both immediate-repeated, $\chi^2(1, N = 90) = 4.94$, $p = .026$, Cramer's V = .23, and delay-repeated participants, $\chi^2(1, N = 60) = 11.88$, $p = .001$, Cramer's V = .45, were more likely to say Yes than immediate-once participants. There was a non-significant tendency for delay-repeated participants

[5] The assumption of homogeneity of variance was violated; therefore the Welch *F*-ratio is reported.

to say Yes more than immediate-repeated participants, Fisher's exact $p = .055$, Cramer's V = .21. Therefore, both repeating evidence with and without a delay increased compliance.

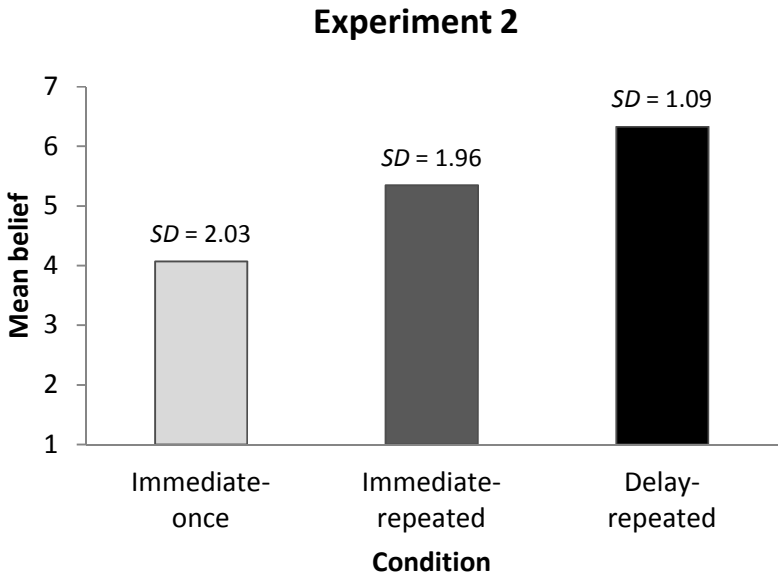

Figure 6. Mean self-reported belief ratings as a function of condition in Experiment 2.

Were participants more likely to falsely believe they cheated when evidence was repeated? A Kruskal-Wallis test revealed that self-reported belief scores differed significantly across conditions, $\chi^2(2, N = 120) = 19.73$, $p = .000$, $\eta^2 = .17$.[6] Follow-up Mann-Whitney *U* tests showed that delay-repeated participants reported higher belief than immediate-repeated participants (average $\text{rank}_{\text{delay-repeated}} = 52.93$ vs. average $\text{rank}_{\text{immediate-repeated}} = 41.78$; $z = -2.08$, $p = .038$). Delay-repeated participants also reported higher belief than immediate-once participants (average $\text{rank}_{\text{delay-repeated}} = 40.18$ vs. average $\text{rank}_{\text{immediate-once}} = 20.82$; $z = -4.46$, p

[6] As with Experiment 1, non-parametric tests were used due to normal distribution and homogeneity of variance violations.

= .000). Immediate-repeated participants reported higher belief than immediate-once participants (average $rank_{immediate\text{-}repeated}$ = 0.93 vs. average $rank_{immediate\text{-}once}$ = 34.63; z = -2.89, p = .004). Thus, participants were more likely to falsely believe they had cheated when they were presented with false evidence more than once, especially with a delay (Figure 6).

An increase in both compliance and false belief with repeated evidence suggests that participants might also be more likely to make up details about how and why they cheated. As in Experiment 1, responses were classified as *No, Partial* or *Full justification* after judges coded for confabulation using Redlich and Goodman's (2003) modified criteria (Table 3). Participants partially justified if they speculated about what might have happened and fully justified if they described how the cheating occurred. Judges agreed on 84.2% (k¼.74) of cases and disputed cases were placed in the more conservative category. Significantly more immediate-repeated, $\chi^2(1, N = 90) = 4.99$, p = .026, Cramer's V = .24, and delay-repeated participants, $\chi^2(1, N = 60) = 13.61$, p = .000, Cramer's V = .48, justified fully than immediate-once participants. Significantly more delay-repeated participants also justified fully than immediate-repeated participants, $\chi^2(1, N = 90) = 4.36$, p = .037, Cramer's V = .22. When collapsed across partial and full justification categories, significantly more delay-repeated participants justified than either immediate-once participants, $\chi^2(1, N = 60) = 11.43$, p = .001, Cramer's V = .44, or immediate-repeated participants, $\chi^2(1, N = 90) = 4.47$, p = .034, Cramer's V = .22. There was a non-significant trend for immediate-repeated participants to justify more than immediate-once participants, $\chi^2(1, N = 90) = 3.45$, p = .063, Cramer's V = .20. When taken together, these justification

statistics suggest that participants were more likely to offer details about how they cheated when repeatedly shown false evidence, and even more so when this evidence was repeated over time.

Table 3. *Judges' justification ratings by condition in Experiment 2.*

	Judges' justification ratings		
	No	Partial	Full
Immediate-once participants	36%	32%	32%
Immediate-repeated participants	30%	30%	40%
Delay-repeated participants	10%	27%	66%

Summary of findings

In summary, presenting false evidence more than once with or without a delay enhanced its effect, presumably because the evidence became more familiar and thus credible (Foster et al., 2012). Repetition with a delay was also more powerful than immediate repetition, which builds substantially on the results from Experiment 1. Indeed, when evidence was repeated with a delay, participants were 20% more likely to justify how they cheated than when evidence was repeated without a delay. Interestingly, Experiment 1 found a null false evidence effect when the evidence was presented immediately after the event. In Experiment 2, this null effect was overcome when the evidence was immediately repeated. The justification ratings of delay-repeated participants in Experiment 2 were also much higher than for late-evidence participants in Experiment 1. Although not

directly tested, this finding suggests that repetition further enhances the delayed effect of false evidence.

Theoretical implications

The results from Experiment 2 suggest that although previous studies have shown the importance of credibility in determining whether or not people believe a source is true (e.g., Nash, Wade, & Brewer, 2009), this role is mediated by both repetition and timing. This study supports previous findings (e.g., Oppenheimer, 2006) suggesting that fluency—mistaken for familiarity—influences peoples' perception of credibility, which in turn acts as an indicator of "truth". Repeating false evidence makes it seem more familiar and thus more credible (e.g. Foster et al., 2012). Repeating it over time increases this effect because people are more likely to turn to external evidence to tell them what happened when their memory of the event has faded. Thus, simply having a sense of déjà vu can make people trust external evidence over their own memory.

The findings suggest that Mazzoni and Kirsch's Metacognitive Model (2002) should refer to a pre-existing belief about familiarity. According to the model, the accusation of cheating should prompt a search in the participant's memory for the event. Because the accusation is false and thus never occurred, the search should be unsuccessful. The participant should then decide that this lack of memory is diagnostic of the event not having occurred and thus not develop a false belief of having cheated. The participant should supposedly only engage in further inferential processes to determine if the event happened, if the event is affected by one of three metacognitive beliefs. Delayed false evidence

corresponds to one of these beliefs, that "memory for an event fades over time". Thus late-evidence participants in Experiment 1 and delayed-repeated participants in Experiment 2 should go on to process the event inferentially and potentially develop a false belief. However, immediate-repeated evidence does not correspond to this belief, or to the other two beliefs that people are likely to experience amnesia for events when they were infants, or that people tend to remember distinctive events. Thus another metacognitive belief must prompt this inferential processing, or else participants should not develop a false belief of having cheated. This metacognitive belief could be for example, "highly familiar events are likely to have occurred", and thus tap into the feeling of fluency that repetition provides.

If the inferential processes described above lead the participant to think that the event is highly likely, this might then lower their memory criterion and induce a false memory. Repetition might revive memories of previously retrieving that memory, which serves to bind different aspects of the event together and thus promote a feeling of remembering (Lindsay, 2008). Repetition might also increase the fluency of the event itself. As Henkel and Coffman (2004) suggest, vivid images become easier to create and—combined with lowered criteria induced by repetition—can lead to not only false confessions, but also cognitive distortions and memory illusions. It is possible that repetition makes the evidence feel fluent and familiar, and thus participants rely on a familiarity heuristic to assign the cheating to memory. The participant may rely on heuristic rather than systematic processing and remember only the vividness and detail as opposed to the cognitive effort that went into forming those imaginings (Henkel & Coffman,

2004). Because vividness and fluency are highly characteristic of perceived rather than imagined events, the mental event would then meet the criterion for being judged as real, and thus participants would develop a false memory of having cheated. This fits with studies which show that exposure to repeated misinformation can lead to false memory (Zaragoza & Mitchell, 1996) and that this effect is exacerbated by increasing vividness, detail and fluency (Mitchell & Zaragoza, 1996). The familiarity of the evidence also likely indicates that the evidence is credible (Foster et al., 2012), and this perceived credibility causes the participant to lower their memory criterion and assign the cheating to memory.

Practical implications

Why does it matter that repeated false evidence increases false belief and justification? In a legal context the findings are important because repetition is used by police throughout suspect interviews. For instance, UK Police must clarify and challenge a suspect's account of an event when necessary, and are allowed to be "persistent" in their line of questioning (ACPO, 2009). Prominent training manuals in the US actively encourage police officers to repeatedly make statements affirming the suspect's guilt and to present suspects with false evidence (e.g., Inbau et al., 2005). However, the findings from Experiment 2 suggest that when false evidence is used, repeatedly presenting it—especially over time—could promote illusory beliefs and false confessions.

The findings also have broader implications in other contexts, such as advertising. Customers are often deceived through advertisements, despite blatant false advertising being illegal in most countries. Food colouring for example, is

sometimes used to make food products appear healthier and more desirable. Taken alongside studies which show that false evidence can lead to changes in behaviour as well as attitude (e.g., Sacchi et al., 2007), this experiment suggests that customers may come to believe false claims and potentially buy products if they are repeatedly presented with deceptive adverts, especially if these adverts are shown over time.

Finally, the results from Experiment 2 may also have implications for everyday experiences. Research shows that thinking about 'what might have been'—or counterfactual thinking—motivates behaviour change and improves performance (Markman, McMullen, & Elizaga, 2008; Morris & Moore, 2000). Perhaps these behavioural changes would be greater if the counterfactual thinking were accompanied by doctored evidence and considered repeatedly over specific delays. Future research should examine this.

Methodological implications

Experiment 2 once again demonstrates the success of the new driving task procedure and shows how it can be used to test other false evidence effects. Importantly, the findings from Experiment 2 are consistent with the key finding of Experiment 1: false evidence becomes more persuasive over time. Experiment 2 also directly tested and were consistent with the tentative data collected in Experiment 1, which suggested that the delay between event and evidence is influential, rather than the delay between event and test.

Interestingly, Experiment 2 indicated that the modality in which the evidence is presented may be of importance. Although not directly comparable,

the early-evidence condition in Experiment 1 and the immediate-once condition in Experiment 2 were identical, except that in the former participants viewed a false video, and in the latter a false photograph. The full justification rates for example are much higher for the video than for the photograph (68% vs. 32%). This observation led to investigating the relative influences of different modalities of false evidence in Experiment 3. For instance, are people more likely to believe fake text or a fake photograph?

To further perfect the driving task methodology, three more changes were made for Experiment 3. First, the number of driving video clips were reduced from 14 to 10 to make the procedure shorter, and thus less laborious. Second, the belief scale was further expanded from 1(strongly disbelieve) to 9(strongly believe) for increased sensitivity. Third, a remember/know/unsure measure was added from Lane and Zaragoza (1995) to provide a direct measure and clearer indication of false memory. As Lindsay (2008) suggests, a variety of measures provide a better insight. Importantly, Mazzoni and Kirsch (2002) point out that "believing" is not the same as "knowing". Belief is simply a judgment that something is true and does not necessarily involve memory. Whereas people understand they have a memory for something that they "know" happened. Therefore both knowing and remembering are essentially indications that people think they have a memory of that event, and provide a different and complementary measure in Experiment 3.

Chapter Summary

Experiment 2 examined what happens when false evidence is shown more than once. The results indicate that repetition increases the power of false evidence, and—in line with the findings from Experiment 1—this effect is exacerbated when the evidence is repeated over time. The experiment clarifies a methodological issue apparent in Experiment 1 and has practical implications for police interrogation and advertising. Importantly, the results are not entirely in line with the Metacognitive Model, and suggest a more prominent role of familiarity in the evaluation process.

Chapter 6: Experiment 3

"But words are more powerful than anything."

—Jennifer Donnelly (1963-Present)

As discussed in previous chapters, studies have shown that false images, videos and narratives can induce people to believe in and develop false memories for both recent and childhood events (e.g., Desjardins & Scoboria, 2007; Nash & Wade, 2009; Wade et al., 2002). However, we do not know whether false images or text are more likely to induce false confessions, beliefs and memories, or what happens when the two modalities are combined. For example, did the modality of the fabricated evidence in Experiment 1 (video) and Experiment 2 (photo) influence its effect? Experiment 3 compares the effects of different modalities directly, using the new driving task method.

Studies have shown that imagination can lead people to develop false autobiographical beliefs and memories about an event—even bizarre events—and that this effect is increased with exposure to false evidence (Mazzoni & Memon, 2003; Nash, Wade, & Lindsay, 2009; Thomas & Loftus, 2002). On the one hand, photographs might help people to visualise the event, and these imaginations might then become confused with real memories because they have similar qualities, such as vividness and high amounts of perceptual detail (Johnson et al., 1993). These rich perceptual details likely "reactivate" similar details from the original event memory, and provide a "cognitive springboard" from which other details and imagined information can be easily generated, thus increasing the

likelihood of source monitoring errors (Johnson, Kahan, & Raye, 1984; Lyle & Johnson, 2006; Weinstein & Nash, 2013). It is also possible that photographs appear more credible than other types information. As discussed in previous chapters, people use credibility as an indicator of truth (e.g., Nash, Wade, & Brewer, 2009) and like repetition, photographs might work by increasing the perceived credibility of the suggestion. Thus if photographs increase both imagination and credibility, they should lead to increased levels of false confessions, belief and memory.

On the other hand, photographs might actually constrain imagination and force people to incorporate certain details. Text might have the opposite effect by giving free reign to the imagination and allowing for a more personally integrated account of events. As Johnson and colleagues (1993) suggest, reading often promotes imagery about the event, which is why people will commit a source monitoring error and think they have seen something that they have actually only read (Belli, Lindsay, Gales, & McCarthy, 1994). Indeed, Garry and Wade (2005) found that participants were significantly more likely to report a false memory of having taken a taken hot air balloon ride when they viewed narratives rather than a doctored photograph of the event. The narrative likely made the event seem familiar because the unconstrained imaginations induced a feeling of being easy on the mind or fluent. This familiarity was then used as a basis to decide that the event had in fact occurred, and thus promoted false memory development (Garry & Wade, 2005; Whittlesea, 1993). Taken together these findings suggest that text might induce more false confessions, beliefs and memories than photographs.

Aims of Experiment 3

In the preceding chapter it was noted that the video used in Experiment 1 appeared to be a more powerful form of false evidence than the photograph used in Experiment 2. However, the experiments are not directly comparable, and comparing data across studies is not advisable because the participants have not been randomly assigned within a controlled experiment. Also, the stimuli in the driving task—the driving clips—are always presented in video format. False video evidence is therefore more similar to the original stimuli than any other type of false evidence. The false video contains the same characteristics and qualities that the original bonus video contained. This makes a comparison between modalities difficult because people are more likely to report false memories when the false evidence has similar qualities to the original information (Lindsay, 2008). Thus, it is possible that participants would be more likely to develop a false memory of cheating when a video is shown rather than a photograph simply because the video contains highly similar details to the original stimuli, rather than because a video is a particularly potent form of evidence. Experiment 3 therefore investigates the effects of modality of false evidence using the new driving task method but without using video evidence.

There were three aims of Experiment 3. The first aim was to investigate the effects of false text and photographs on false confessions, beliefs and memory. Although studies such as Garry and Wade's (2005) have shown the impact of these modalities on false autobiographical memory, they have not specifically studied their impact in the context of false confessions. In the previous chapter it was noted that in the stressful situation of being falsely accused, participants

might take short-cuts when evaluating information. For instance, participants may use the vividness and familiarity of mental content to make a source decision, and neglect to consider any associated cognitive effort (Henkel & Coffman, 2004). Because of this "short-cut" in processing, photographs might induce the same—or possibly higher—level of false confessions than text. The cognitive effort required to overcome the constraints placed on the imagination by the photographic details might be ignored.

The second aim of Experiment 3 was to investigate the effects of using both modalities—text and photos—together, because previous research has only focused on the power of using one modality. It is possible that there might be additive effects, such that using text and photographs in combination might induce more false confessions than either used alone. It is also possible that the free reign provided by the text is constrained by the photograph, leaving participants confused about what "fits" with their event memory. For instance, after viewing the text, a participant might think, "Yes, I must have clicked on red when I was looking at that other car". However, when presented with the photograph, the same participant might think, "No, I wasn't looking at the car at that point... this makes no sense at all". This might then prompt them to engage in more effortful source monitoring and thus, discredit the false evidence and reject the accusation that they cheated.

The final aim was to once again refine the new driving task procedure and to make any necessary amendments to improve it. The design of Experiment 3 needed to differ from Experiments 1 and 2 in one important way. Whereas the previous two experiments showed false evidence based on the actual mouse clicks

that participants had made during the bonus clip, the evidence in this experiment was based on a single "made-up" inserted click. This added experimental control and enabled the text and photograph to depict the same information from the same point in time. The piloting and creation of these materials is discussed in the Methods section below.

EXPERIMENT 3

Method

Participants

One hundred and twenty adults, 17-59 years-old, were recruited from the University of Warwick to take part in a driving test. Participants took part voluntarily, for course credit, or for £3 and were randomly allocated to a 2 (Photograph: Yes vs. No) × 2 (Text: Yes vs. No) between-participants design. This allowed for four groups: *photograph* (n = 30, M = 19.80 years, SD = 2.12), *text* (n = 30, M = 21.00 years, SD = 6.27), *both* (n = 30, M = 22.20 years, SD = 8.55) and *control* (n = 30, M = 20.17 years, SD = 3.72).[7]

Materials and Procedure

The procedure used was a modified version of Experiment 2. The procedure was streamlined to last only 20 minutes, including only 11 video clips instead of 15, ranging from 32-95 s in length (M = 52.55 s, SD = 19.43). Participants were seated at a computer and instructed to click the mouse when they thought they saw a hazard on screen. Clicking when a red light was present was classed as cheating, which would result in their disqualification. Participants were told that they would be given the chance to win £50 on the bonus clip, which was always Clip 3. Immediately after this clip all participants were falsely informed that they had cheated. Participants then continued with the rest of the test for approximately 6.5 mins. All false evidence was then presented at the end of the test for 25 s.

[7] An additional 36 participants who were tested but cheated on the bonus clip, and one who indicated suspicion were excluded.

There were four conditions. Photograph participants were shown a false photograph of the supposed cheating, which depicted a single frame of the video with a red light and red flag indicating a click (see Figure 7). Text participants were shown the equivalent information in the form of a 54 word computer record. To create the computer record, eight judges viewed the photograph used in the photograph condition and answered five questions: "What type of area was

Figure 7. False photograph used in Experiment 3, depicting the supposed cheating.

this photograph taken in?" "What season of the year was this photograph taken in?" "What is the weather like in this photograph?" "What is the likely speed limit of this area?" "What is happening in the photograph?" They were also asked about anything else they could infer from looking at the photo. Judges listed 80 details in total. Details reported by four or more (50%) judges were used to construct the text. The final version of the computer record was as follows:

Number of clicks on red light: 1; Clip number: 3

Type of area: Residential; Time of year: Autumn/Winter; Weather: Dry

Speed limit: 30mph

Frame details: After junction/traffic lights; Parked vehicles on left; 2 parked white vans on right; Car breaking and oncoming car in distance

Frame features: Cyclist road markings; Waste bins on pavement

Both-participants were shown the photograph and text simultaneously. The layout of this was counterbalanced, such that either the text or photograph was at the top of the screen. Control participants were told they had cheated but were not shown any false evidence. The total number of clicks made during the bonus clip did not significantly differ across conditions, $F(3, 62.58) = .74$, $p = 0.53$.[8]

The experimenter, who was blind to participants' conditions, then interviewed participants about their feedback on their software. The experimenter read aloud 6 statements, one of which was critical "I cheated on the bonus clip". Participants were asked to agree or disagree with the statement, which served as a measure of compliance. Participants were then asked to rate how much they believed the statement to be true on a 9-point scale from 1(strongly disbelieve) to 9(strongly believe), as a measure of belief. Participants were also asked to describe and/or explain what happened, which provided a measure of justification. If participants agreed they had cheated on the bonus clip, they were asked to choose from the options remember/know/unsure. The experimenter clarified these

[8] The assumption of homogeneity of variance was violated; therefore the Welch *F*-ratio is reported.

terms when necessary using the following definitions adapted from Lane and Zaragoza (1995):

> Remembering means that you can consciously recall some aspect or aspects of what happened, or what you experienced at the time.
>
> Knowing means that although you believe you clicked on a red light, you cannot consciously recall what happened or what you experienced at the time.
>
> Unsure means you are not sure at all about whether your memory corresponds to either remembering or knowing.

Results and Discussion

Did modality of false evidence influence false confession rates? Loglinear analysis revealed a significant Photograph x Text x Compliance interaction, LR $\chi^2(1) = 7.62$, $p = .006$ (see Figure 8). Breaking down this interaction shows that photograph participants were no more likely to comply than control participants, LR $\chi^2(1) = 1.85$, $p = .176$. However, text participants were more likely to comply than either control participants, LR $\chi^2(1) = 21.67$, $p < .001$, or photograph participants, LR $\chi^2(1) = 12.33$ $p < .001$. Photographs appeared to lessen the impact of text when shown alongside, such that text participants were significantly more likely to comply than both-participants, LR $\chi^2(1) = 5.83$, $p = .016$.

Did text also induce more false belief than a photograph? Self-reported belief scores differed significantly across conditions according to a Kruskal-Wallis test, $\chi^2(3, N = 120) = 19.20$, $p < .001$, $\eta^2 = .16$.[9] Follow-up Mann-Whitney *U* tests revealed that text participants reported higher belief than either control participants (average $\text{rank}_{\text{text}} = 38.53$ vs. average $\text{rank}_{\text{control}} = 22.47$; $z = -3.69$, p

[9] As with Experiments 1 and 2, violations of normal distribution and homogeneity of variance required the use of non-parametric tests.

<.001) or photograph participants (average $rank_{text}$ = 36.70 vs. average $rank_{photograph}$ = 24.30; z = -2.87, p = .004). Once again, photograph participants

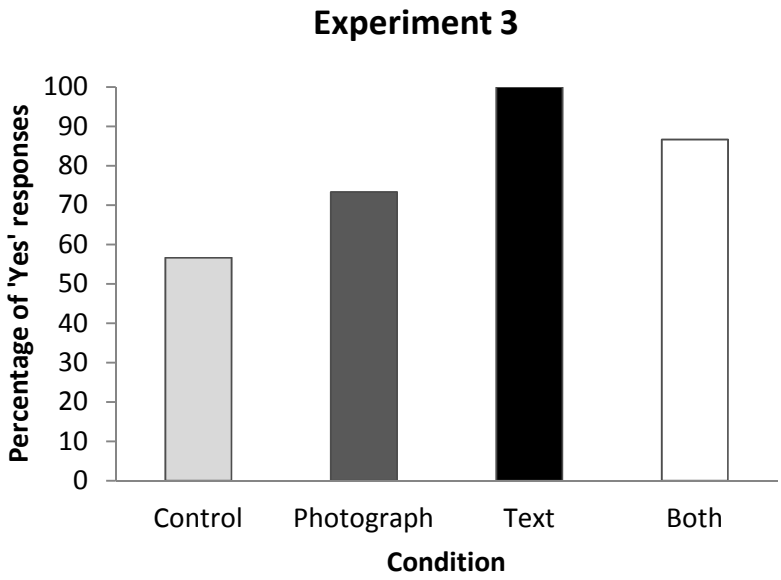

Figure 8. Percentage of participants complying as a function of condition in Experiment 3.

did not significantly differ to control participants (average $rank_{photograph}$ = 32.53 vs. average $rank_{control}$ = 28.47; z = -.91, p = .361). Both-participants did not significantly differ to text participants (average $rank_{both}$ = 30.25 vs. average $rank_{text}$ = 30.75; z = -.12, p = .902). However, both-participants reported higher belief than either control (average $rank_{both}$ = 37.13 vs. average $rank_{control}$ = 23.87; z = -3.05, p = .002) or photograph participants (average $rank_{both}$ = 35.87 vs. average $rank_{photograph}$ = 25.13; z = -2.48, p = .013). Thus, the presence of text always appeared to increase false belief and this effect was not lessened by the addition of a photograph. This was also the case for "remembering". Only 6.7% of control and 6.7% photograph participants said they could explicitly remember

having cheated, compared to 23.3% of text participants and 26.7% of both-participants. Showing participants text significantly increased the rate of remembering relative to other responses, LR $\chi^2(1) = 7.98$, $p = .005$, and this effect was not diminished by adding a photograph, LR $\chi^2(1) = .09$, $p = .766$.

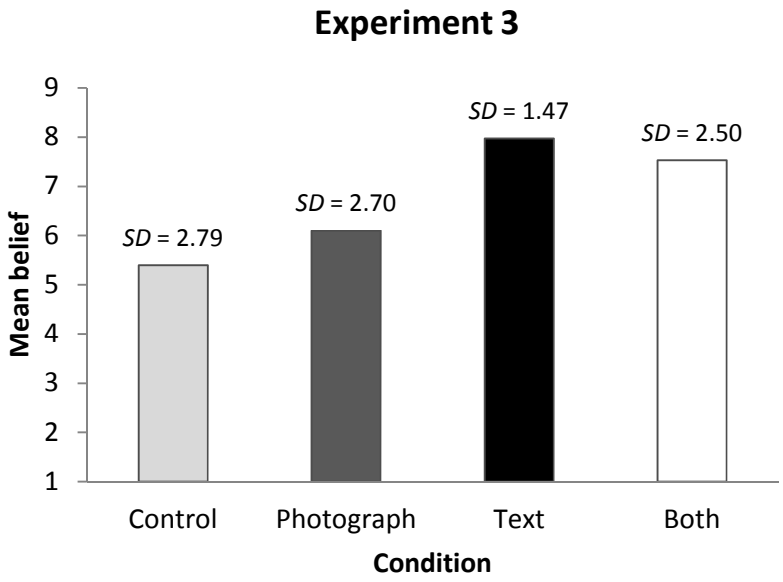

Figure 9. Mean self-reported belief ratings as a function of condition in Experiment 3.

If participants were more likely to say they "remembered" cheating when shown false text over a photograph, were they also more likely to justify how the cheating occurred (Table 4)? Judges coded for confabulation using Redlich and Goodman's (2003) modified criteria, and these classifications were then relabelled *No, Partial* or *Full justification* as in Experiments 1 and 2. Judges agreed on 85.0% (k¼ = .77) of cases and disputed cases were always placed in the more conservative category. Participants partially justified by postulating about what

might have happened and fully justified by outlining how the cheating happened. For full justifications, loglinear analysis did not reveal a three-way interaction, LR $\chi^2(1) = 1.93$, $p = .164$. However, there was a significant main effect of Text, LR $\chi^2(1) = 14.47$, $p < .001$. Smaller loglinear analyses revealed that text participants were significantly more likely to be judged as full justifying than either control, LR $\chi^2(1) = 13.08$, $p < .001$, or photograph participants, LR $\chi^2(1) = 4.40$ $p = .036$. Showing a photograph alongside text did not reduce the power of text to elicit full justifications, LR $\chi^2(1) = .07$, $p = .795$. When partial and full justifications were collapsed, a three-way Photograph x Text x Justification interaction was found, LR $\chi^2(1) = 5.43$, $p = .020$. Once again this interaction was driven by text participants justifying significantly more than either control, LR $\chi^2(1) = 25.77$, $p < .001$, or photograph participants, LR $\chi^2(1) = 13.57$ $p < .001$. However, there was also a marginal difference between text participants and both-participants, LR $\chi^2(1) = 3.21$, $p = .073$, such that adding a photograph to text tended to decrease justification compared to text by itself.

Table 4. *Judges' justification ratings by condition in Experiment 3.*

	Judges' justification ratings		
	No	Partial	Full
Control participants	60%	27%	13%
Photograph participants	40%	30%	30%
Text participants	3%	40%	57%
Both-participants	17%	30%	53%

Summary of findings

In summary, false text is more powerful than a false photograph in inducing false confessions, false belief and justification. Interestingly, the powerful effect of text becomes less potent in some cases when shown alongside a photograph. Participants were more likely to say they had cheated and there was a tendency to justify more often when shown text in isolation rather than text and a photograph together. However, participants were no more likely to believe they had cheated or to remember cheating when shown only text rather than both. Together these data suggest that the addition of a photograph to text may have had a subtle effect, such that it discredited the convincing text evidence and left participants confused as to what really happened.

Interestingly, Experiment 3 reveals that a false photograph shown once after a brief delay is no more persuasive than misinformation by itself. As Experiment 2 indicated, photographs appeared to be less influential than video (Experiment 1). It could be that photographs need a longer delay to have an effect, such as in the delay-repeated condition in Experiment 2. This experiment shows that text, however, is still powerful after only a brief delay of 6.5 mins.

Theoretical implications

The results from Experiment 3 fit with the findings of Experiment 2. Experiment 2 suggested that people are more likely believe something is true—and potentially develop a false memory—if the false evidence they see feels highly familiar. According to Mazzoni and Kirsch's (2002) Metacognitive Model,

this feeling of familiarity in Experiment 2 should not have influenced participants' beliefs or memory unless they had a pre-existing belief such as "if a mental image feels familiar then it's likely to have happened". If participants did hold such a belief, then they would go on to consider the highly familiar image of cheating against a set of criteria, and thus decide whether or not they cheated on the driving task. Experiment 3 once again suggests that familiarity is an important factor in developing a false belief or memory. Instead of repetition, it is the modality of the false evidence that influences how familiar the mental image of cheating becomes.

Why was text a much more powerful modality of false evidence than a photograph? Research has shown that text allows people to draw inferences from what they read (Brewer, 1977), which are often vividly remembered and recalled with confidence (Chan & McDermott, 2006). It is highly likely that people are able to make these inferences because the text promotes imagination. In this experiment, text presumably allowed the participant to freely imagine the cheating scenario and build up a coherent mental picture of how events could have occurred. Imagination in itself promoted false memory because it involves mental rehearsal and thus repetition. This rehearsal then made the text easy to access, familiar and fluent. Interestingly, false recall of fluent stimuli is often accompanied by a vivid memory and strong recollective experience (e.g., Roediger & McDermott, 1995). Presumably then, a fluent and familiar image of the cheating might have promoted not only false belief, but also false memory too. Indeed, as Mazzoni and Kirsch (2002) point out, mental content is more likely to be judged as a memory when a factor makes it come to mind very easily. The clear and vivid mental picture of clicking on a red light that easily came to mind

likely became confused with the original memory, which contained similar identifying characteristics (Johnson et al., 1993). The participant then committed a source monitoring error by judging that the false cheating occurred. The increased familiarity and fluency of the image might also have made the text evidence appear more credible (Foster et al., 2012). The perception of increased credibility might then have lowered the memory criterion and thus cheating was erroneously attributed to memory.

Previous studies have also shown that photographs can be a particularly potent form of suggestion, leading people to falsely believe in events that never happened (e.g., Henkel, 2011a; Nash & Wade, 2009). For instance, seeing a photograph of a completed action can lead people to claim that they themselves completed this action, with repetition increasing the number of false claims (Henkel, 2011a; Henkel & Carbuto, 2008). This effect has been termed the *photograph inflation effect* and is thought to occur because photographs provide a "cognitive springboard" from which mental content—images, thoughts, feelings—can easily be generated (e.g., Strange, Garry, Bernstein, & Lindsay, 2011). This mental content has similar characteristics to the actual event memory and thus becomes easily confused.

However, in this experiment, false photographs were not persuasive and did not promote source confusion above a control condition. This could be because—as mentioned above—photographs are not as persuasive as other forms of false evidence. Indeed, looking at the early-evidence condition in Experiment 1 and immediate-once condition in Experiment 2, suggests that—as well as text—videos are also more powerful in inducing false belief and justification than

photographs. False photographs may need to be repeated, or shown after more than a delay of just 6.5 mins to be effective. Indeed, existing false photograph studies have used longer delays or discussed a distant childhood event over several sessions (e.g., Henkel, 2011a; Strange et al., 2007; Wade et al., 2002). As demonstrated in Experiment 1, people are more likely to turn to evidence presented after a delay to tell them what happened, presumably because their original event memory has faded. Thus, in these previous studies, participants have used the photographs to build up a hypothetical mental picture consistent with the false scenario, and this mental content has not conflicted with their dim recollection of events.

The lack of persuasive power of a photograph is best explained using the Self-Memory System (SMS) Model discussed in Chapter 2 (Conway & Pleydell-Pearce, 2000; Conway et al., 2004). In the driving task scenario, photographs likely constrained the imagination with the explicit details they contained. The participant likely detected discrepancies between the photograph and their still relatively vivid memory of the event, and thus questioned the credibility of the photograph. The inconsistent details may have prompted a narrow search of mental content for associated material, most likely resulting in rejection of the false suggestion. Whereas text promoted a feeling of fluency and a broad search for material, photographs may have actually promoted a feeling of "disfluency". The salient unfamiliar details that were provided in the photograph—a mouse click at an unfamiliar point in time on a red light—may have resulted in a feeling of cognitive dissonance. Cognitive dissonance is a negative state of conflict that can occur after making a decision, which we try to avoid by gathering supportive

information and ignoring unsupportive information (Festinger, 1957). Indeed, unfamiliar details in a photograph may have more of a fluency-impeding effect than in text because the participant has less free reign of the imagination to view the event in their own manner (Hessen-Kayfitz & Scoboria, 2012).

If photographs promote such "disfluency" though, why did photograph participants not show lower false belief and memory rates than control participants? Photograph participants still appeared to believe the evidence to the same extent that they believed the accusation of cheating by itself. Indeed, this was even the case when a photograph was presented immediately after the accusation for immediate-once participants in Experiment 2. This finding is somewhat surprising given the presence of contradictory details in the photograph. It is possible that the feeling of disfluency prompted by the photograph was just enough to make participants sceptical about the credibility of the evidence, but not enough to discredit the accusation completely. This is likely because, although the photograph contained inconsistent details, it was based on a true frame taken from the actual original video, and thus also contained many consistent details. These true details likely induced false belief in the accusation similarly to text, but this belief was then dampened by the other contradictory details. This was even more likely for immediate-once participants in Experiment 2, because the false photograph included an actual click they had made at a given point in time, and thus the evidence appeared even more "true". The idea that the evidence in these experiments is partially true is discussed in more detail in Chapter 10.

It is an interesting point that salient and potentially "disfluent" details present in the photograph are also presumably present in a false video. If this is the case, why was a false video so persuasive in Experiment 1? Are videos perceived to be outwardly more credible than other types of evidence to the extent that they lower the belief and/or memory criterion considerably more? Is it because the video in Experiment 1 depicted actual clicks that the participant had made as opposed to an inserted click in the photograph in this study, and thus was more "true"? Or perhaps, is it because there was a slightly longer delay (9 mins vs. 6.5 mins) in Experiment 1 than in Experiment 2, and thus participants' memories had had longer to fade? Further research should address these questions and directly compare the effects of false video evidence with other modalities.

What happened then when both false text and a false photograph were presented together? It is possible that participants did not devote much attention to the photograph when it was shown alongside text. In the brief time that both pieces of evidence were presented, participants' cognitive resources could have been overloaded with information (e.g., Basil, 1994). Thus participants may only have focused on the highly influential text and ignored the photograph. Indeed, text is generally thought to require more effortful cognitive processing than pictorial information and thus demands more attentional resources (e.g., Hsieh & Chen, 2011). However, this would likely have resulted in a null effect of adding a photograph to text. Contrary to this outcome, in some cases, simultaneously presenting a photograph with text actually appeared to decrease the persuasive power of text. Thus, it is more likely that some of the photographic details created a conflict for the participant because they were inconsistent with the imagined

inferences of the text. Indeed, previous research suggests that showing a photograph alongside text should only promote false recall if the photograph is consistent with the inferences likely drawn from the text (Henkel, 2011b). The inconsistency between inferences and photographic details, combined with the constrained imagination, likely caused participants to question the credibility of both pieces of evidence, and thus engage is more effortful source monitoring.

This explanation also fits with decision making research on the *dilution effect* (e.g., Nisbett, Zukier, & Lemley, 1981; Shanteau, 1975). When people see two independent pieces of evidence—one strong and one weak—they tend to illogically judge the overall strength of the evidence to be somewhere in between the two, instead of adding the strength of the two supportive pieces together. When both pieces of evidence were presented together in this experiment, the evidence strength might have been "averaged", such that the weak photograph diluted the strong text evidence.

Interestingly, The SMS model does not predict the pattern of results shown by both-evidence participants. According to the SMS, participants who saw both a photograph and text should have falsely confessed and developed false memories less than those shown just a photograph or just text. This is because the details generated from text that were consistent with the original event memory would have conflicted with the potentially inconsistent details present in the photograph. This should have resulted in the narrowest search of all for associated details, and thus the lowest rates of false confessions and memories. However, in Experiment 3, control and photograph participants showed the lowest rates. This suggests that text is so powerful and convincing that the photograph only has a minimal

diminishing effect when shown alongside. Indeed, this fits with the results that depict only a subtle effect of adding a photograph to text.

Importantly, Experiment 3 is the first to have investigated these combined effects of false text and a false photograph. What happens though when different types of evidence are shown separately over time? Do people focus on the first piece of evidence they see and then disregard any subsequent pieces? Or do people sum up all of the evidence and form an average opinion? Experiment 4 addresses these questions. Experiment 3 also raises questions in relation to the other experiments in this thesis. For instance, Experiment 2 showed that presenting a photograph more than once, especially over time, increased its power substantially and made it a persuasive form of false evidence. If text is already persuasive when shown only once as demonstrated in Experiment 3, might it be especially persuasive when shown repeatedly? This question was also investigated in Experiment 4.

Practical implications

In terms of policing, it is important to look at the different forms in which evidence is presented. As discussed previously, the use of false evidence during questioning is allowed and encouraged in some US jurisdictions. Experiment 3 suggests that written information—such as a false eyewitness statement—might be an especially powerful form false evidence, which could elevate the number of false confessions and thus wrongful convictions. The findings from Experiments 1 and 2 also suggest that false videos may be more powerful than a false photograph. This has implications for CCTV evidence, which may be presented to

suspects during interrogation. As suggested above, future research should compare these modalities directly using a different methodology in which the original stimuli are not presented in video format.

The results of Experiment 3 could be extended to the presentation of true rather than false evidence. Even in countries where false evidence is prohibited during interviewing, true evidence is occasionally presented. For instance, a witness might be shown a photograph of the crime scene to help them mentally time travel back to the event and thus aid their recall. This in itself can be dangerous as Lindsay et al. (2004) demonstrated, when they showed participants true photographs of an event along with other suggestive influences and in doing so lead participants to develop false autobiographical memories. Experiment 3 suggests that the effect could be even more potent when a true piece of text—such as an eyewitness statement—is used instead of a photograph.

Similarly to Experiment 2, the findings from Experiment 3 also have wider implications. For instance, advertisements that reference autobiographical events have been found to induce false memories in people (Braun et al., 2002). Interestingly, Braun and Loftus (1998) found that pictures and words in advertisements induce an equal number of false memories, and that pictures tend to evoke more "remembering" responses. This suggests that the traditional reliance on pictorial information—such as in perfume advertisements (Tuna & Freitas, 2012)—could be dangerous if the advertisements are misleading or autobiographical. Yet the results from Experiment 3 suggest that written information could in fact be more dangerous than pictorial information. Indeed, there may be an interaction effect of the modality in which the information is

presented and that in which the participant is tested (Braun-LaTour et al., 2004). This interaction should be the participant of future research.

In the clinical domain, trauma therapists sometimes attempt to "trigger" childhood memories during therapy (Poole, Lindsay, Memon, & Bull, 1995). One way of doing this is to peruse childhood photographs. Experiment 2 suggests that this might be dangerous if the photographs are seen repeatedly over time. Taken together with Garry and Wade's (2005) findings regarding false narratives, Experiment 3 suggests that looking at other written materials from childhood—diary entries or letters for instance—might be even more dangerous than photographs. If misleading, these materials could potentially lead to the development of false childhood beliefs and memories.

Chapter Summary

Experiment 3 looked at whether the false evidence effect is influenced by modality. Perhaps counter-intuitively, the findings show that text is more persuasive than a photograph. The results also suggest that when both types are shown together, the photograph might diminish some of the text's power. Like Experiment 2, this experiment also emphasises the role of familiarity in belief and memory attribution, and has potential implications for police practice, advertising and therapy. The results are also partially explained by the SMS model.

Chapter 7: Experiment 4

"It is folly to put the plough in front of the oxen."

—François Rabelais (1494-1553)

Are there order effects when different types of false evidence are shown? What happens when different modalities are presented over time rather than simultaneously as in Experiment 3? Previous studies (e.g., Desjardins & Scoboria, 2007; Garry & Wade, 2005) have shown the persuasive power of text in developing false childhood memories, and Experiment 3 demonstrated that this effect carries over into a confession scenario. However, we do not know whether text is always persuasive, or whether its power is influenced by the timing of its presentation relative to other less powerful modalities. To this end, Experiment 4 used the driving task procedure to vary the order in which participants viewed a false photograph or written statement.

Research suggests that the first piece of evidence people are presented with should be more influential than successive pieces of evidence. People tend to lock onto the first piece of information they see and use it to "anchor" the rest of their judgments (Tversky & Kahneman, 1974). This *anchoring effect* occurs in a wide variety of contexts. For instance, both mock-jurors and legal experts have been found to place more weight in the evidence they receive first (e.g., Chapman & Bornstein, 1996; Englich, Mussweiler, & Strack, 2006). When presented with new information, people re-evaluate the original information and update their judgments. However, the judgments still remain relatively close to the original

anchor (e.g., Epley & Gilovich, 2006). Simply thinking about the anchor and comparing the new information to it, makes the anchor itself feel more "fluent" and easy on the mind. Any information associated and consistent with the anchor then also becomes more accessible. This ultimately produces a *confirmation bias* in which people accept information consistent with their anchor and reject or ignore new inconsistent information (see Nickerson, 1998 for a review).

The results from Experiment 3—combined with anchoring research—suggest that false confession rates would be higher when the first piece of evidence participants see is false text rather than a photograph. Participants are more willing to accept the accusation that they cheated when they are shown text, because text likely prompts participants to freely imagine the cheating scenario with details included that are consistent with the original event memory. However, a photograph may contain details that conflict with the event memory and thus places constraints on the imagination. Thus, if text is more persuasive, and the first piece of evidence is more influential, participants who see text first should falsely confess, believe and remember the cheating more than those who see a photograph first.

However, Experiment 3 also showed that the power of text to induce compliance and justification was lessened somewhat when shown alongside a photograph. Presumably, the inconsistent details present in the photograph not only impeded fluency for the photograph itself, but also for the text it was presented with. Thus although people may place more emphasis on the first piece of information they see, if a second piece of evidence contains inconsistent and conflicting details it may cause participants to question the credibility of all of the

evidence. If the credibility of the evidence is lowered, it might not meet the belief and memory criteria and thus the SMF and Mazzoni and Kirsch's (2002) Metacognitive Model would predict that the cheating suggestion would be rejected. Therefore, showing false text first might not result in high false confession, belief and memory rates if a photograph is then presented afterwards.

Aims of Experiment 4

The primary aim of Experiment 4 was to investigate anchoring effects of different modalities of fabricated evidence on false confessions, beliefs and memory. Previous studies have investigated anchoring effects in the development of false memories but not in the context of false confessions. For instance, Wade, Garry, Nash and Harper (2010) investigated the anchoring effects of narratives and photographs in the development of false childhood memories. Participants viewed either a photograph first or a narrative first of a fake childhood hot air balloon ride. Wade et al. found that participants who saw the narrative first were more likely to develop false memories of the event than those who saw the photograph first. However, because of the two group between-participants design of the study, it was impossible to tell if the second piece of evidence had any influence on memory. Thus the secondary aim of Experiment 4 was to address this issue by employing a 2 (First evidence: Photograph vs. Text) × 2 (Second evidence: Photograph vs. Text) between-participants design.

EXPERIMENT 4

Method

Participants

One hundred and twenty adults from the University of Warwick, 17-54 years old, took part for course credit or £3. Participants were randomly assigned to the cells of a 2 (First evidence: Photograph vs. Text) × 2 (Second evidence: Photograph vs. Text) between-participants design. This resulted in four experimental groups: *photo-photo* (n = 30, M = 21.27 years, SD = 2.90), *text-text* (n = 30, M = 19.87 years, SD = 2.90), *photo-text* (n = 30, M = 20.70 years, SD = 3.15) and *text-photo* (n = 30, M = 22.30 years, SD = 8.43).[10]

Materials and Procedure

Experiment 4 used the same procedure and stimuli as Experiment 3 except for two changes. First, participants viewed only eight video clips instead of 11, ranging from 32-95 s in length (M = 50.78 s, SD = 20.01). This streamlined the procedure and allowed for minimal delay between evidence presentations to avoid ceiling effects. Second, all participants viewed one piece of evidence immediately after the bonus clip and a second piece of evidence at the end of the test, approximately 5 mins later. Each piece of evidence was shown for 12.5 secs. Photo-photo participants viewed the false photograph from Experiment 3 both times. Text-text participants viewed the false computer record statement from Experiment 3 both times. Photo-text participants viewed the photograph first and

[10] An additional 21 participants who were tested but cheated on the bonus clip, and one who indicated suspicion were excluded.

the computer record second. Text-photo participants viewed the computer record first and the photograph second.

Results and Discussion

Did the order of the evidence influence how often participants falsely confessed to cheating? Loglinear analysis did not reveal a significant First evidence x Second evidence x Compliance interaction, LR $\chi^2(1) = 2.15$, $p = .143$. However, there was a significant main effect of First evidence, LR $\chi^2(1) = 11.64$, $p = .001$. Text-text participants were significantly more likely to comply than either photo-photo participants, LR $\chi^2(1) = 5.76$, $p = .016$, or photo-text participants, LR $\chi^2(1) = 11.85$, $p = .001$, but not text-photo participants, LR $\chi^2(1) = 1.12$, $p = .290$. Text-photo participants were also significantly more likely to comply than photo-text participants, LR $\chi^2(1) = 6.26$, $p = .012$, but not photo-photo participants, LR $\chi^2(1) = 1.97$, $p = .161$. There was no significant difference in compliance between photo-photo participants and photo-text participants, LR $\chi^2(1) = 1.28$, $p = .258$. Thus in general—but not always—if participants saw text first they were more likely to comply than if they saw a photograph first, regardless of what type of evidence was shown second. Indeed, 93% of text-first participants complied compared to 70% of photo-first participants.

Order of evidence also affected self-reported false belief rates, according to a Kruskal-Wallis test, $\chi^2(3, N = 120) = 19.71$, $p < .001$, $\eta^2 = .17$.[11] Follow-up Mann-Whitney *U* tests revealed that text-first participants reported higher belief than photo-first participants (average $\text{rank}_{\text{text-first}} = 73.33$ vs. average $\text{rank}_{\text{photo-first}} = 47.68$; $z = -4.11$, $p < .001$). The persuasive effect of showing text first occurred

[11] As with Experiments 1, 2 and 3, the ordinal data violated the normal distribution and thus required the use of non-parametric tests.

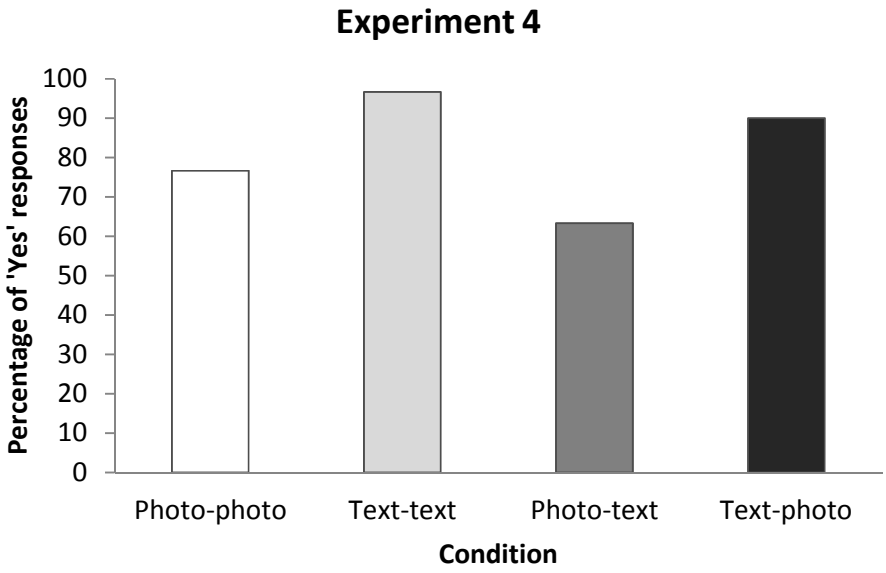

Figure 10. Percentage of participants complying as a function of condition in Experiment 4.

regardless of what type of evidence was presented second. Text-text participants indicated higher belief than either photo-photo (average $rank_{text\text{-}text}$ = 35.18 vs. average $rank_{photo\text{-}photo}$ = 25.82; z = -2.13, p = .033) or photo-text participants (average $rank_{text\text{-}text}$ = 38.67 vs. average $rank_{photo\text{-}text}$ = 22.33; z = -3.69, p < .001) but did not differ from text-photo participants (average $rank_{text\text{-}text}$ = 30.30 vs. average $rank_{text\text{-}photo}$ = 30.70; z = -.092, p = .926). Text-photo participants also reported higher belief than photo-photo (average $rank_{text\text{-}photo}$ = 35.20 vs. average $rank_{photo\text{-}photo}$ = 25.80; z = -2.13, p = .033) or photo-text participants (average $rank_{text\text{-}photo}$ = 38.60 vs. average $rank_{photo\text{-}text}$ = 22.40; z = -3.64, p < .001). Photo-photo participants also reported marginally higher belief than photo-text participants (average $rank_{photo\text{-}photo}$ = 34.47 vs. average $rank_{photo\text{-}text}$ = 26.53; z = -1.78, p = .075).

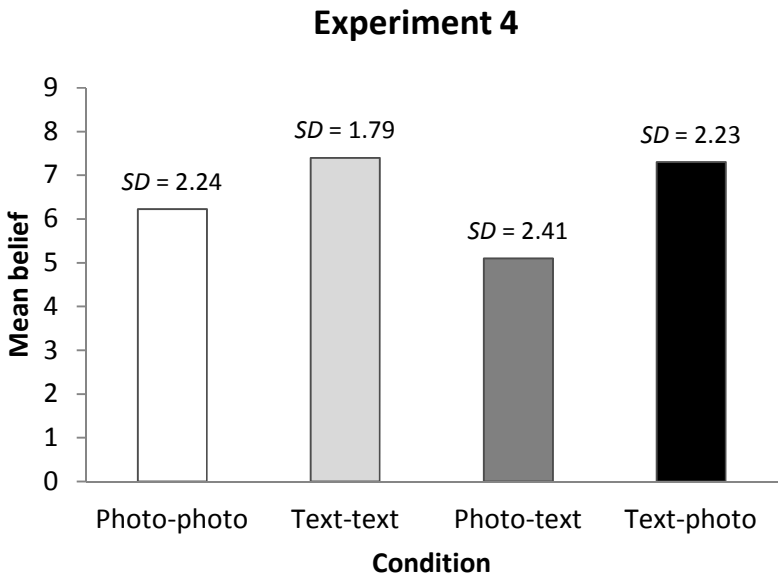

Figure 11. Mean self-reported belief ratings as a function of condition in Experiment 4.

There were no significant main effects or three-way interaction found for participants claiming to "remember" the cheating relative to other responses (*p*s ≥ .706). However, collapsing across "remembering" and "knowing" categories revealed a significant main effect of First evidence, LR $\chi^2(1) = 25.90$, $p < .001$. This is important because as Mazzoni and Kirsch (2002) point out, both remember and know responses suggest that the participant thinks they hold a memory for the event. Text-text participants were more likely to remember/know than either photo-photo participants, LR $\chi^2(1) = 19.01$, $p < .001$, or photo-text participants, LR $\chi^2(1) = 9.93$, $p = .002$, but not more so than text-photo participants, LR $\chi^2(1) = .07$, $p = .787$. Text-photo participants were also more likely to say they remembered or knew they cheated than photo-photo participants, LR $\chi^2(1) = 16.89$, $p < .001$, or photo-text participants, LR $\chi^2(1) = 8.35$, $p = .004$. There was no difference

between photo-photo and photo-text participants, LR $\chi^2(1) = 1.69$, $p = .193$. Thus text-first participants were more likely to report remembering or knowing they cheated than photo-first participants irrespective of the second piece of evidence they were shown.

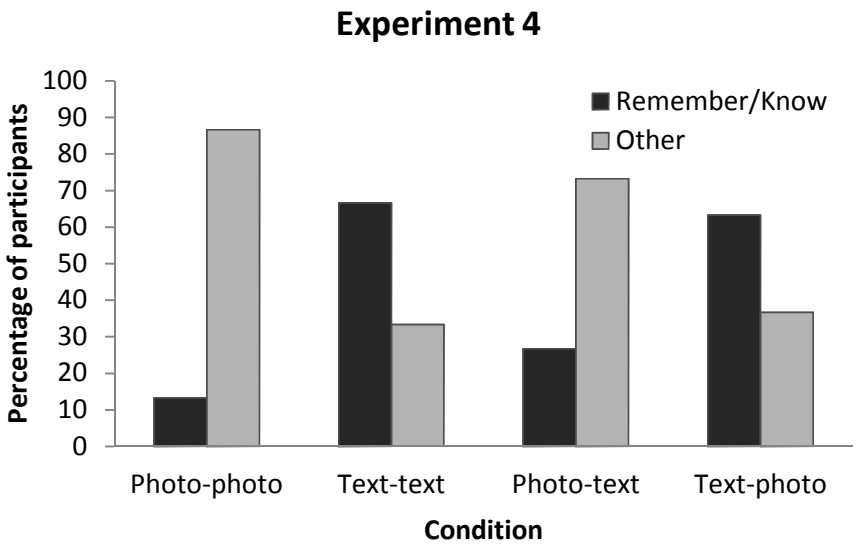

Figure 12. Percentage of participants reporting remembering/knowing they cheated versus other responses by condition in Experiment 4.

The remember/know responses matched justification data, providing an overall indication of memory. Judges used Redlich and Goodman's (2003) modified criteria to code for confabulation, and as in previous experiments, these classifications were relabelled *No, Partial* or *Full justification* (see Table 5 and Figure 13). Judges agreed on 80.8% (k¼ = .71) of cases and disputed cases were always placed in the more conservative category. Participants partially justified by postulating about what might have happened and fully justified by outlining how

the cheating happened. Loglinear analysis did not reveal a First evidence x Second evidence x Full justification interaction, LR $\chi^2(1) = 0.25$, $p = .275$. However, there was a main effect of First evidence, LR $\chi^2(1) = 7.90$, $p = .005$. Text-text participants were significantly more likely to fully justify details about how the cheating occurred than either photo-photo or photo-text participants, both LR $\chi^2(1) = 4.49$, $p = .034$. Text-text and text-photo participants did not fully justify significantly differently, LR $\chi^2(1) = .07$, $p = .787$, but text-photo participants marginally fully justified more than either photo-photo or photo-text participants, both LR $\chi^2(1) = 3.44$, $p = .064$. There were no significant full justification differences between photo-photo and photo-text participants, LR $\chi^2(1) = 0.00$, $p = 1.000$.

Table 5. *Judges' justification ratings by condition in Experiment 4.*

	Judges' justification ratings		
	No	Partial	Full
Photo-photo participants	43%	43%	13%
Text-text participants	10%	53%	37%
Photo-text participants	57%	30%	13%
Text-photo participants	20%	47%	33%

The same pattern of results was observed when partial and full justification categories were collapsed. There was no significant First evidence x Second evidence x Justification interaction, LR $\chi^2(1) = 2.22$, $p = .136$, but there

was a main effect of First evidence, LR $\chi^2(1) = 17.44$, $p < .001$. Text-text participants were significantly more likely to justify than photo-photo, LR $\chi^2(1)$

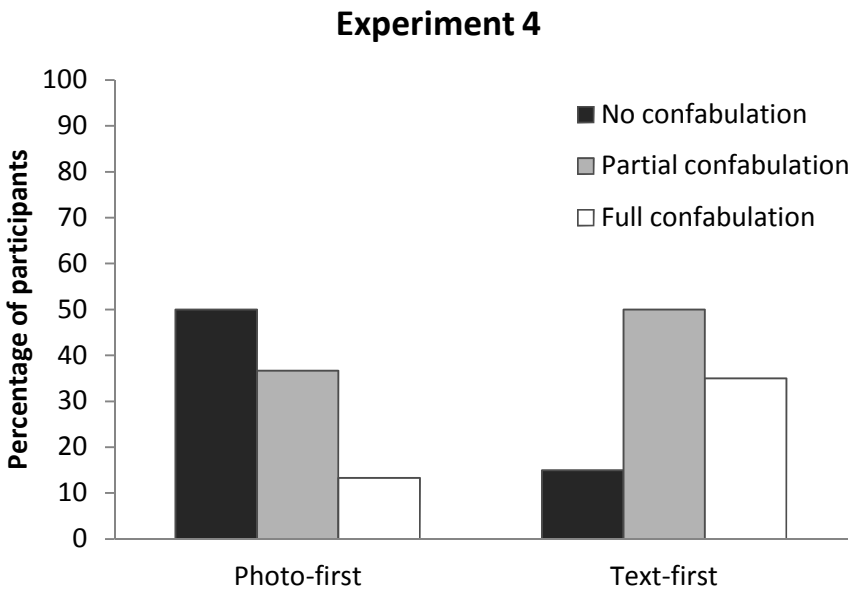

Figure 13. Percentage of participants indicating no, partial or full justification when shown either a photograph first or text first in Experiment 4.

= 9.03, $p = .003$, or photo-text participants, LR $\chi^2(1) = 15.82$, $p < .001$, but not more than text-photo participants, LR $\chi^2(1) = 1.20$, $p = .274$. Text-photo participants also justified significantly more than photo-photo, LR $\chi^2(1) = 3.84$, $p = .050$, or photo-text participants, LR $\chi^2(1) = 8.80$, $p = .003$. Photo-photo and photo-text participants did not justify significantly differently, LR $\chi^2(1) = 1.07$, $p = .301$. Thus the justification results follow the same pattern as the other measures, such that participants were more likely to justify when they saw text first rather than a photograph first, and this did not depend on what type of evidence was shown second.

Summary of findings

In sum, showing false text first is almost always more persuasive than showing a photograph first. This occurs regardless of what type of evidence is presented second, and is robust across compliance, belief, remember/know and justification measures. This suggests an anchoring effect, such that we decide on whether to accept the false evidence as being true based on only the first piece of evidence we encounter. After a brief delay, the second piece of evidence presented appears to be discarded. The results of Experiment 4 are in line with Experiment 3, which already demonstrated the superior effects of false text over a false photograph. Experiment 4 shows that the powerful effect of text holds when it is presented first.

At first glance, the anchoring effects of Experiment 4 appear to conflict with the repetition results found in Experiment 2. According to Experiment 4, if people make a decision based solely on the first piece of evidence, then seeing a second piece of evidence should have no effect. However, this is contrary to the influential delayed-repeated evidence effects found in Experiment 2. It is highly likely that there is an interaction between repetition, modality and anchoring. People may be more persuaded by repeated information—regardless of its modality—compared to information only presented once. However, this repetition effect is heightened when the first piece of evidence is presented in a powerful modality (e.g., text) because more weight is placed on the initial piece of evidence when making the judgment.

Theoretical implications

Theoretically, the results from Experiment 4 extend the findings from Experiment 3. Experiment 3 demonstrated the powerful effect of text in inducing false confessions relative to a photograph. It was suggested that text promotes unconstrained imagination of the cheating scenario, thus prompting a broad search for associated mental content for the event, and often resulting in acceptance of the cheating suggestion. Whereas, a photograph was thought to restrict the imagination due to the potentially conflicting perceptual details, resulting in a narrow search and most likely rejection of the cheating scenario. The findings from Experiment 4 once again demonstrate the superiority of text over a photograph in inducing false confessions, presumably for the same theoretical reasons.

However, in Experiment 4, text was only powerful when presented first. This demonstrates a clear anchoring effect, such that participants decided whether or not they cheated based solely on the first piece of evidence they saw. Because Experiment 4 employed a 2x2 design, it was possible to see that the second piece of evidence had no effect on judgment. This extends Wade et al.'s (2010) findings and emphasises the importance of the initial anchor. Participants rejected any inconsistent information that was presented after the first piece of evidence. Even details in the false photograph—which likely sometimes conflicted with participants' event memories—were rejected if participants initially viewed text, and the inconsistency appeared to have no influence on the perceived credibility of the overall evidence or source. Experiment 4 suggests that participants show conformation bias when presented with false evidence, such that they only accept successive information if it is consistent with the first piece of evidence they see.

The results from Experiment 4—when different modalities are shown separately over time—do not match the results from the both-evidence condition in Experiment 3, in which different modalities were shown simultaneously. In Experiment 3, text was always highly persuasive, even when shown alongside a photograph. There was only a very subtle diminishing effect of the photograph. Experiment 4 shows that a photograph is highly detrimental to false memory development if shown first, yet it has no effect at all if shown second. Once again, this demonstrates the importance of anchoring effects when evidence is shown over time. This also contradicts the dilution effect—as discussed in the previous chapter—(e.g., Nisbett et al., 1981; Shanteau, 1975). Presenting a weak piece of evidence (a photograph) after a strong piece of evidence (text) had no dilution effect. Anchoring effects overshadowed dilution effects, such that when there was a delay between the two pieces of evidence, participants only focussed on the initial strong piece of evidence.

Similar to the both-evidence results from Experiment 3, the findings from Experiment 4 do not completely fit with the SMS model of memory (Conway & Pleydell-Pearce, 2000; Conway et al., 2004). The model suggests that if participants see both text and a photograph, this prompts a very narrow search in memory for mental content associated with the cheating, and thus likely results in more rejections of the accusation than if participants see solely text or a photograph. This is supposedly because the participant is forced to reconcile the consistent and fluent details generated from the text with the inconsistent and "disfluent" details present in the photograph (e.g., Hessen-Kayfitz & Scoboria, 2012). According to this model, the photo-text and text-photo conditions should

therefore have resulted in the highest rejection rates. However, this was not the case, because photo-text was just as persuasive as photo-photo and text-photo was just as persuasive as text-text. Indeed, the results from Experiment 4 suggest that when shown separately over time, ultimately anchoring effects decide whether or not false evidence is accepted as being true, rather than conflict because of consistent and inconsistent details. Importantly, Experiment 4 is the first study to have fully investigated how anchoring moderates the false evidence effect.

Practical implications

Experiment 3 suggested that showing suspects written information, such as a false eyewitness statement, could be especially dangerous during police interrogation in terms of increasing false confession rates and ultimately wrongful conviction. Experiment 4 further highlights this important finding. The results from Experiment 4 also demonstrate the complexity of showing suspects multiple pieces of false evidence. Experiment 4 brings together the findings of all of the experiments in this thesis, to show that timing, repetition, modality and ordering of false evidence all interact to mediate and moderate the persuasive false evidence effect. Not only does it matter when and how many times the evidence is shown (Experiments 1 and 2), it also matters what modality and the order that the modalities are presented in (Experiments 3 and 4). In terms of policing, it is important to look at the different forms in which evidence is presented. For instance, false confession rates could potentially be much higher if a false eyewitness statement is presented before a false photograph of the crime scene than if presented the other way around. Like the results on Experiment 3, the

findings of Experiment 4 could also be extended to the use of true evidence in interrogation. This is a matter for future research.

Experiment 4 also extends the wider implications from Experiment 3. The results of Experiment 3 suggested that misleading messages in advertisements might be more persuasive if presented as text rather than images. This was contrary to previous findings and the current heavy reliance on images in advertising (e.g., Braun & Loftus, 1998; Tuna & Freitas, 2012). The results of Experiment 4 further suggest that text might only be persuasive if presented before any images. Potential implications of these findings are discussed in Chapter 10.

Chapter Summary

Experiment 4 investigated whether the power of text still holds when it is presented either before or after a photograph. The results demonstrate a strong anchoring effect, such that people only use the first piece of evidence they see when making a belief and memory judgment. Once again, the results highlight the importance of fluency and have implications for police practice and advertising. Like Experiment 3, the findings are also partially in line with the SMS model. The findings bring together the results from all of the experiments in this thesis, and suggest interactions between different factors.

Chapter 8: Digital editing questionnaire

"Yes, of course we do post-production corrections on our images. Photoshopping is an industry standard."

—Danzinger, editer of *Self* magazine (2009, as cited in Hartmann, 2009)

How aware are people of digital editing? Can their perceptions of digital materials be changed? Can science alter the way in which people approach digital materials? With the increasing use of digital editing software in the media, governments have recently begun tightening regulations for its use in advertising. For instance, in 2011 the UK Advertising Standards Authority banned a cosmetics advertisement featuring Julia Roberts because extreme airbrushing promoted an unrealistic picture of flawless skin (Zhang, 2011). In the US, the National Press Photographers Association (NPPA, 2013) released a code of ethics warning journalists not to "manipulate images or add or alter sound in any way that can mislead viewers or misrepresent participants". Studies show that editors themselves are reasonably tolerant of minor digital alterations, and that this tolerance is higher for magazine as opposed to news editors (Reaves, 1987, 1991). But how do the readers of these media—the general public—view digital editing, and is it possible to educate people about digital editing through experiments?

A handful of studies have investigated public perceptions of digital materials and have found mixed results regarding awareness of—and attitudes towards—digital editing. For example, 40% of a community sample in one study did not know that digital alterations had been used in the industry for more than

20 years (Huang, 2001). This finding suggests a lack of awareness of the extent to which digital editing has previously been used. However, participants in this study were generally concerned about the use of digital alterations in current media and 65% suggested guidelines for the practice. Suggestions included minimal and moral use of doctored images in appropriate contexts, and importantly participants wanted to be notified when images had been altered.

Yet simply knowing that a photograph has been altered does not change peoples' perceptions of how credible a news photograph is (Greer & Gosen, 2002). When participants were shown one of five versions of a news photograph, their perceived credibility of the photograph only decreased as the level of alteration increased. Thus, people are likely to be tolerant of minor digital alterations in the media, but major alterations need justification. However, being exposed to such major alterations in this study did not affect participants' attitudes towards digital media overall.

It appears that even when participants have strong negative attitudes towards digital editing, they are unlikely to actively oppose the practice. Reaves, Hitchon, Park and Yun (2004) conducted a study in which young females viewed both unaltered and digitally altered photographs of models featured in the popular fashion magazine *Vogue,* before answering a series of questions regarding the photographs and digital alteration in general. Unlike in Huang's (2001) study, participants in this study displayed a high awareness of digital editing, although they reported rarely editing digital media themselves. Participants disapproved of using the practice to make models appear thinner, and deemed the practice to be a manipulative selling tool which was dishonest, unfair and unethical. Despite these

seemingly strong views, participants were reluctant to take action against the digital editing. Thus it appears that neither awareness nor strong opinions can influence peoples' acceptance and behavioural intentions towards digital editing in the media.

Aims of the questionnaire

Participants from Experiment 3 completed a questionnaire on their perceptions of digital editing. The purpose of this study was twofold. First, the questionnaire collected data on peoples' use and awareness of digital editing software. The extent to which participants actually use software themselves has been largely neglected in the psychological literature, and data on general awareness is conflicting. Second, the questionnaire was used to investigate how the experimental manipulation in Experiment 3—that is, the modality in which evidence is presented—might influence participants' opinions and intentions regarding digital materials. Reaves et al.'s (2004) study of fashion model photographs showed that participants intended to be more careful in future when viewing magazine photographs following their study. Other studies have shown that being deceived during an experiment changes peoples' expectations and behaviour in future experiments despite being debriefed (Jamison, Karlan, & Schechter, 2008; Krupat & Garonzik, 1994). If exposure to digital alterations and deception during an experiment can both affect behavioural intentions, then deceiving people about the authenticity of digital materials—as in Experiment 3—might affect participants views' and behavioural intentions towards digital materials in the future. For instance, being exposed to a doctored photograph in

Experiment 3 might lead participants to intend to consider digital materials more carefully in the future. The current study tested this hypothesis.

DIGITAL EDITING QUESTIONNAIRE

Method

Participants

The participants were the same 120 participants, 17-59 years old, who took part in Experiment 3. Participants were randomly allocated to one of four groups: *control* (n = 30, M = 20.17 years, SD = 3.72), *photograph* (n = 30, M = 19.80 years, SD = 2.12), *text* (n = 30, M = 21.00 years, SD = 6.27), and *both* (n = 30, M = 22.20 years, SD = 8.55).

Materials and Procedure

Participants filled out a 9-item questionnaire regarding their experiences with digital editing (Appendix B). A further eight participants piloted an original version of the questionnaire which included one additional question that asked participants what materials they thought might be digitally altered extensively. Participants had a tendency to reiterate example items that had been given in previous questions (e.g., photograph in a magazine) and thus the results were biased. This question was removed because it did not yield any useful data.

At the end of Experiment 3 all participants received the same debriefing: They were informed that they had been deceived because they had not cheated in the task but were misinformed they had done so. The experimenter was blind to the condition participants were in and thus equally described all experimental conditions to the participants and explained the reason for investigating these conditions. The experimenter explained that deception was necessary to study the development of false beliefs and confessions, and gave the participants the opportunity to ask any questions.

After this debriefing, participants were asked to complete a digital editing questionnaire that took less than 5 mins to complete on average, but participants were given unlimited time. Participants were asked to answer nine questions with regards to digital editing, which was defined as "anything from electronically cropping a photograph, to professionally cutting and editing videos". The first five questions asked about participants' use of digital editing software: how often they used it and what for, what they edited and with what software, and what level of expertise they held. The next two questions investigated general perceptions of digital editing: how often participants considered that materials might have been digitally altered and how much they trusted materials they thought might have been altered. The final two questions investigated the effects of participation in Experiment 3 on participants' perceptions of digital editing. Participants were asked if the experiment had increased their awareness of digital editing and whether they would be more careful in evaluating digital materials in the future. The questions required a mixture of forced choice, 5 point Likert scale, and open responses.

Results and Discussion

Analyses were broken down into three main categories based on the question groupings in the questionnaire. First, use of digital editing software was analysed descriptively. Second, the correlation between consideration and trust of digital materials was analysed. Third, whether the experimental manipulation in Experiment 3 affected participants' behavioural intentions towards digital materials in the future was investigated.

Use of digital editing software

Figure 14 shows that approximately one third of participants (35%) reported that they 'sometimes' used digital editing software and approximately half of all participants reported using it rarely or never (52%). Frequent usage was uncommon, with only one participant reporting use all the time. Participants reported using the software mostly for leisure purposes (45%) and for editing photographs (65%). Out of participants who said they used digital software, over three quarters (78%) rated themselves at novice level. Although these results suggest that participants' use of digital editing software is not particularly high, the results are still higher than those obtained by Reaves et al. (2004), who found that participants reported scarce use of digital alteration technology. Indeed, in their study participants reported a mean use of 2.21 (*SD* = 1.82) on a scale ranging from 1(never) to 7(frequently). It is likely that in the eight years between the two

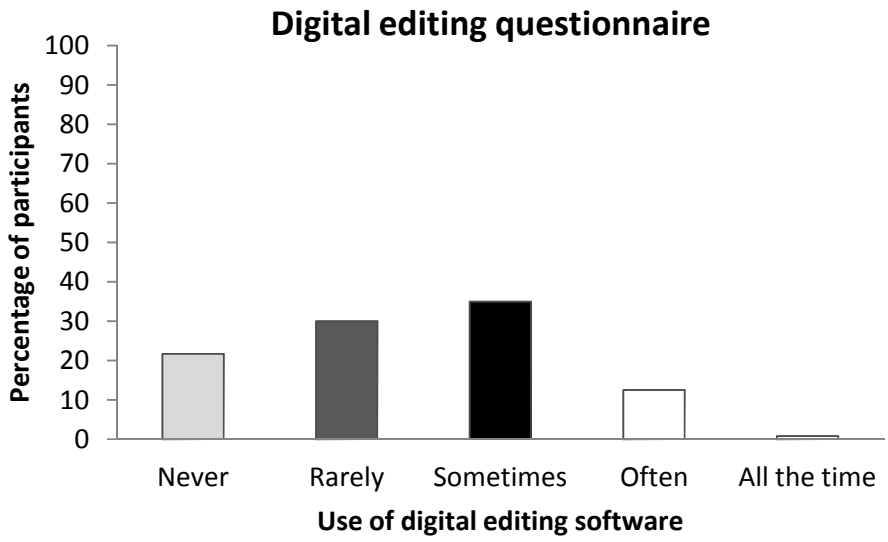

Figure 14. Percentage of participants self-reported use of digital editing software in the digital editing questionnaire.

surveys, digital technology has become even more prevalent and thus use of digital editing has increased. Indeed, there are several lines of argument for students entering university now being "digital natives" (e.g., Kennedy, Judd, Churchward, Gray, & Krause, 2008).

Consideration and trust of digital materials

Despite only moderate use of digital editing software themselves, the majority of participants reported considering that materials may have been digitally altered either sometimes or often (73%). Approximately half of participants (51%) reported rarely trusting materials they thought had been altered. However a large proportion (40%) of participants reported sometimes trusting these materials, despite perceiving that they had likely been edited in some way. A Spearman's rho analysis revealed a significant positive correlation between consideration of materials being altered and trust in materials perceived to have been altered, $rs[120] = .20$, $p = .028$. However, squaring the correlation coefficients indicated a very weak correlation, such that only 4% of the variance was explained by this relationship. Although weak, this correlation suggests that the more participants reported they considered if a material had been digitally altered, the more likely they were to generally trust digital materials. This fits with Greer and Gosen's (2002) finding that experience with imaging software increased tolerance of digital manipulation of news photography. At first glance these findings appears counter-intuitive: Surely the more often we scrutinise a material for having been altered in some way and have experience with alterations ourselves, the more sceptical we become and the less likely were are to trust

digital materials? However, Greer and Gosen (2002) suggest that ever increasing accessibility of digital technology might be leading us to assume that all digital materials have been altered. Thus it is possible that this finding reflects a general acceptance of digitally altered materials. This is in line with Huang's (2001) research which showed that the more participants used a computer and were aware that most photographs in the news and magazines are altered, the more accepting they were of digital alterations. Future studies should investigate the relationship between consideration, trust and acceptance of materials.

Future awareness and evaluation of digital materials

Did actually being exposed to—or simply being educated about—digital alteration in Experiment 3 affect participants' behavioural intentions towards digital materials in the future? Photograph, $\chi^2(1, N = 60) = 5.46$, $p = .020$, Cramer's V = .30, and both-participants, $\chi^2(1, N = 60) = 11.88$, $p = .001$, Cramer's V = .45, were more likely than control participants to say Yes when asked if the experiment had made them more aware of digital editing. There were no other significant differences between conditions for rated awareness, all $ps \geq .091$. The same pattern was observed for participants' intentions to scrutinise future materials following the experiment. Photograph, $\chi^2(1, N = 60) = 8.09$, $p = .004$, Cramer's V = .37, and both-participants, $\chi^2(1, N = 60) = 14.60$, $p < .001$, Cramer's V = .50, were more likely to say Yes than control participants when asked if they would evaluate digital materials more carefully in the future. No other differences were significant, $ps \geq .089$, except that both-participants intended to evaluate future materials significantly more than text participants, $\chi^2(1, N = 60) = 5.46$, $p =$

.020, Cramer's V = .30. Thus participants who had viewed a false photograph during the test—regardless of whether or not they had also viewed text—reported that they would be more aware of digital editing and evaluate digital materials more carefully in the future. This suggests that participation in Experiment 3 educated participants about a practice they already reported to frequently consider. It is possible that photographs rather than text had this effect because photographs have a "concrete believability" that words do not have, and thus altered photographs have more of an impact on trust (Lowrey, 1998, as cited in Greer & Gosen, 2002). Interestingly, this finding was only true for participants who viewed digitally altered photographs themselves. Simply being told about doctored photographs—as in the text condition—did not increase intentions to be more aware of, and evaluate, digital materials more carefully in the future.

Chapter Summary

The questionnaire reported in this chapter examined peoples' perceptions of, and attitudes towards, digital editing. Despite all participants reporting considering that digital materials may have been altered—and 78% of participants reporting using digital software themselves—almost half of all participants still reported trusting materials they perceived to have been altered. This worrying finding suggests that people are relatively accepting of doctored materials, and mirrors Reaves et al.'s (2004) concerns that digital editing in magazines is often deemed "trivial" despite research, for instance, linking thin models in magazines with eating disorders (e.g., Thomsen, McCoy, & Williams, 2001). Indeed, one pilot participant in the current study appeared to sum up this view. In response to

the question, "Do you trust materials that you think might have been altered extensively?", the participant answered, "Depends on the alteration - if it is just cosmetic, there is no need to be paranoid". However, despite these concerns, both Reaves et al. (2004) and this survey show that people can be educated about digital alterations and in doing so their intentions towards the way that they consider these materials in the future can be altered.

Chapter 9: Theoretical implications

"Everything must be taken into account. If the fact will not fit the theory, let the theory go."

—Agatha Christie (1890-1976)

The four experiments in this thesis have shown that the persuasive effect of fabricated evidence is moderated by several different factors. Moreover, it is not just the strength of the effect that is altered, rather at least in terms of timing and modality, it is whether or not fabricated evidence has any effect at all over and above misinformation. This is the most surprisingly finding of this research programme and extends current understanding of the fabricated evidence effect. Despite previous studies showing that false evidence can induce high false confession, belief and confabulation rates (e.g., Horselenberg et al., 2006; Kassin & Kiechel, 1996; Nash & Wade, 2009), this thesis shows that this effect is dependent on how exactly the evidence is presented. This chapter focuses on the possible theoretical implications of this thesis. The first part of this chapter summarizes the potential theoretical findings from each experiment. The second part outlines how these findings might be used to refine and integrate current theories of autobiographical false belief and memory.

Part I: Theoretical findings

Experiment 1

The findings from Experiment 1 demonstrate that the false evidence effect is completely mediated by the timing of the presentation of the evidence. When false evidence is presented immediately after the event, it is no more powerful than misinformation alone. Whereas, when it is presented after a brief delay, false evidence is highly persuasive and induces both false belief and justification.

Experiment 1 highlights the importance of metacognition and how pre-existing beliefs can influence belief and memory formation. The results fit with Mazzoni and Kirsch's (2002) Metacognitive Model described in Chapter 2. The accusation of cheating likely prompts all participants to search for a memory of having cheated, which is inevitably unsuccessful because the event is false. For early-evidence participants, this lack of memory is diagnostic of the event having not occurred, and thus they reject the accusation as false. However, late-evidence participants are likely prompted to engage in further inferential processing because they hold the belief that "memory fades over time". Thus, not having a memory of cheating a short while ago does not necessarily mean it did not happen. Participants may then be convinced by the evidence and come to believe they cheated. Although false memory was not measured in Experiment 1, it is possible that participants may have also developed a false memory, because—in line with the SMF (e.g., Johnson et al., 1993; Lindsay, 2008)—the perceptual details contained in the video promote imagery, and the highly vivid and plausible image of cheating could become easily confused with the true memory.

At first, this experiment does not appear to support the mechanism of credibility. The actual source of the evidence—the computer—does not vary in any of the experiments presented in this thesis. In Experiments 1 and 2, the

evidence itself that is presented is always the same, and thus should not be any more credible in a certain condition. However, in Experiment 1, it is possible that evidence presented immediately after the event appeared less credible than evidence presented late. Early-evidence participants might have detected discrepancies between their recent—and relatively strong—memory of the event and the false evidence, and this might have decreased the perceived credibility of the evidence. However, the memory of the event might have faded for late-evidence participants by the time they were presented with the false evidence, and thus they perceived it to be more credible. This concept also fits with the SMS model of memory (e.g., Conway & Pleydell-Pearce, 2000; Conway et al., 2004). The inconsistencies between the event memory and false evidence detected by the early-evidence participants, could have caused a conflict within the working-self and led to a narrow search of memory. Whereas, late-evidence would have likely triggered a broader and more focussed search for associated mental content.

Experiment 2

Experiment 2 shows that the false evidence effect is moderated by how many times the evidence is presented. Repeating evidence increases false belief and justification, and this effect is exacerbated when the evidence is repeated over time.

Unlike Experiment 1, Experiment 2 highlights the importance of a familiarity mechanism and thus fits with the Memory Attribution Approach (e.g., Jacoby, Kelley, et al., 1989). Repeating the evidence presumably made it feel more "fluent" and easy on the mind. Fluency is taken as an indicator of

familiarity, and—in line with the SMF—familiarity is used as a cue to source, such that mental content that is familiar is more likely to be judged as true rather than imagined.

Although there is no direct evidence supporting a credibility mechanism, it is possible that if something is familiar, this also makes it seem more credible (e.g., Foster et al., 2012). Thus repeated evidence felt more fluent, more familiar and potentially more credible and convincing than evidence only presented once. The powerful effect of repetition was exacerbated when the evidence was repeated over time because this likely tapped into participants' pre-existing belief that "memory fades over time". Repetition also likely promoted imagery by aiding rehearsal of the mental content through imagination.

The findings of Experiment 2 suggest that the Metacognitive Model needs to include a pre-existing belief relating to familiarity. According to the model, the lack of memory of the cheating should be diagnostic that the event did not occur. However, the results suggest that this is not the case, and that actually repetition prompted participants to engage in inferential processing. It is likely therefore that people hold a belief such as "familiar content is likely to be true" and thus they go on to evaluate the information more carefully. It is also possible that repetition prompts people to bypass systematic processing altogether and simply rely on a familiarity heuristic. Fluency aids imagination, and the imagined information becomes highly detailed and vivid and thus automatically fulfils the belief criterion and is accepted as being true.

Experiment 2 says little about the SMS model. Presumably however, the increased fluency that comes from repetition, prompts a broader search for mental

content, and thus is more likely to result in a false memory than evidence only presented once.

Experiment 3

Experiment 3 demonstrates that the modality of false evidence mediates its effect. After a brief delay, a false photograph is no more persuasive than misinformation by itself. However, false text induces high levels of false belief and justification.

Like Experiment 2, the findings of Experiment 3 appear to support a familiarity mechanism, as described in the Memory Attribution Approach. Text likely provides the imagination with free reign to generate details consistent with the cheating scenario. For instance, a participant might think, "I must have clicked when I was distracted by that person walking along the pavement just as the car started to brake in front". This prompts fluent and easy rehearsal of the now familiar mental content. Unlike Experiment 2 however, Experiment 3 has potential implications for the SMS model. The "consistent" details generated from the text—such as the person walking along the street—likely prompt a broad, yet focused, search for mental content associated with the cheating. Whereas, a photograph might also contain vivid perceptual details that conflict with the original event memory. For instance, a participant might think, "I'm sure the light wasn't red when the car went over those big white road markings". This might induce a feeling of "disfluency" and cognitive dissonance, which results in a narrow search for mental content, and thus likely rejection of the false evidence and cheating accusation.

However, the results do not fully support the SMS model. The effect of text was sometimes diminished when presented alongside the photograph. The disfluency of the photograph likely lowered the perceived credibility of the evidence as a whole. However, against the predictions of the SMS model, this both-evidence condition did not result in the lowest rates of false belief and memory because of a highly narrow search. Thus, Experiment 3 shows the ultimate power of text to overcome conflict within the working-self.

Again, like Experiment 2, Experiment 3 also suggests that the Metacognitive Model should include a pre-existing belief related to familiarity that prompts inferential processing.

Experiment 4

Experiment 4 shows that the fabricated evidence effect is moderated by the order in which the evidence is presented. The results highlight the importance of an anchoring mechanism, such that people base their judgment on only the first piece of evidence they see and ignore the second piece. Thus, presenting text first results in higher belief and memory rates than presenting a photograph first, regardless of what type of evidence comes next.

The findings once again show the superiority of text as discussed for Experiment 3. However, text was only superior when shown first and was ignored when shown second. Thus Experiment 4 demonstrates that ultimately anchoring moderates the effect of modality.

Like the both-evidence condition in Experiment 3, Experiment 4 cannot be explained by the SMS model. Conditions in which participants viewed both text

and a photograph did not necessarily result in the lowest rates of false belief and memory, rather it only mattered which piece of evidence was presented first. This demonstrates the power of the anchoring effect, such that is overrides feelings of inconsistency and cognitive dissonance.

Summary and conclusions

The experiments presented in this thesis have potential implications for theories of false belief and memory. Experiments 2-4 suggest that the Memory Attribution Approach (Jacoby, Kelley, et al., 1989; Whittlesea, 1993) is right to emphasise the importance of a familiarity mechanism. However, Experiment 1 suggests that this emphasis on familiarity should not overshadow other important mechanisms, such as credibility and the vividness of imagery. Thus, as posited by the SMF (Johnson et al., 1993; Lindsay, 2008), familiarity should be viewed as one of several cues to source. The findings of Experiment 1 fit closely with the Metacognitive Model (Mazzoni & Kirsch, 2002). The results of Experiments 2-4 however, suggest that the model needs to include a pre-existing belief relating to familiarity that prompts inferential processing. The results of Experiments 3-4 support the SMS model (Conway & Pleydell-Pearce, 2000; Conway et al., 2004) in highlighting the role of inconsistent details in prompting cognitive dissonance. However, the results suggest that the importance of inconsistency is over-emphasised, and that other factors—such as modality and anchoring—have the ability to override these effects.

There are also two other interesting points to note. First, it is not necessarily the credibility of the source—in this case, the computer—that matters,

rather the perceived credibility of the evidence itself. Second, the SMS model posits that only false memories consistent with the current goals of the working-self can be induced in the laboratory. At first glance, it seems unlikely that participants in these experiments held a potential goal of "owning up to cheating" in much the same way that suspects in real life would surely never hold a goal of falsely confessing. However, there are several documented cases in which people have made self-deprecating coerced-internalised confessions (see Kassin, 1997). One explanation could be that people can be persuaded to believe evidence inconsistent with their goals if it is powerful enough. However, an alternative explanation could be that false confessors are simply being consistent with other goals, such as being helpful and trying to please the interrogator.

Part II: A modified Metacognitive Model

The above theoretical findings of this thesis suggest a distinction between autobiographical belief and potential indicators or precursors of memory—(e.g., justification and remember/know measures)—and thus clearly link with Mazzoni and Kirsch's (2002) Metacognitive Model. However, the results do not fully support all aspects of the model and some findings are better explained using other theories that were discussed in Chapter 2. The second part of this chapter posits modifications to the current Metacognitive Model by incorporating important aspects of these other theories, and by using the results from Experiments 1-4. It is important to note here that the experiments comprising this thesis only suggest these modifications to theory; they do not necessarily directly

test or demonstrate these mechanisms. These suggestions are drawn from possible interpretations of the findings from Experiments 1-4, and require empirical testing and further research.

How the current Metacognitive Model could be refined

As discussed in Chapter 2, the Metacognitive Model builds on the SMF, and thus is based on the same principles. Importantly however, the Metacognitive Model distinguishes between the formation of autobiographical beliefs and memories. The findings from this thesis are consistent with this distinction between belief and memory, and thus this section keeps them as separate entities. For instance, 73% of participants who viewed false text in Experiment 3 either believed or fully believed they cheated. However, only 23% of text participants claimed that they "remembered" the cheating. Thus it is important to investigate how false beliefs and memories are formed in isolation and then how a belief can be converted into a memory.

Although it does not necessarily make this distinction, Jacoby's Memory Attribution Approach is important and needs to be considered, because it highlights the importance of familiarity, which was demonstrated in Experiments 2-4. The approach is closely linked with the SMF, and suggests that people misattribute an imagined event to memory when the imagined event is highly fluent and familiar. However, the experiments in this thesis—particularly Experiment 1—suggest that familiarity is not the only cue to source. Thus, the Memory Attribution Approach needs to integrate with the Metacognitive Model, such that familiarity is included, but alongside others cues as well.

Interestingly, research suggests that the dual process signal detection account of recognition (Wixted & Stretch, 2004) can also be used to understand how familiarity might influence recollection (Chan & McDermott, 2006). Studies have shown that people often falsely remember inferences from statements, and that the phenomenological experiences accompanying these false memories are highly similar to those for true memories (Chan & McDermott, 2006; McDermott & Chan, 2006). The authors suggest that this is because although the processes of recollection and familiarity may be separate, both entities are then combined to provide one coherent indicator of memory strength. Thus even without a strong feeling of recollection, a strong sense of familiarity—perhaps from repetition (Experiment 2) or a particularly fluent modality (Experiment 3)—might then result in a person "remembering" an event. This is also potentially why remember/know judgments are not as independent of each other as they may seem. This idea is returned to in the next chapter.

The results of Experiments 3 and 4 were not fully explained by the Metacognitive Model. The findings appeared to fit more closely with the SMS model, which highlights the importance of consistent vs. inconsistent details. For this reason, aspects of the SMS model may inform the Metacognitive Model. Initially, this appears to be problematic, because the SMS is a very different model of memory to the Metacognitive Model and does not cover the development of autobiographical belief. However, this thesis suggests—somewhat controversially—that the SMS model could actually be applied to belief rather than memory, and then a false belief converted into a false memory through the mechanisms suggested in the Metacognitive Model. For instance, if

the prevalence of consistent and fluent details generated by text only leads to the development of false memories—and then following on from this, beliefs—why did half of all text participants in Experiment 3 develop a false belief and not a memory? Presumably, the consistent and fluent details must have led to increased false belief—and then in line with the Metacognitive Model—this belief was in some cases converted into a false memory.

One of the most important elements that the SMS model emphasises—and the Metacognitive Model does not—is the influence of the "self" in autobiographical memory. The experiments in this thesis, alongside the results of other studies, suggest that the self perhaps needs to be considered in models of autobiographical memory. For instance, in a series of experiments, participants rated the extent to which they could "relive" certain memories, and how accurate they believed them to be (Rubin, Schrauf, & Greenberg, 2003; Rubin & Siegler, 2004). Visual and auditory imagery, emotions and the coherence of the story predicted the level of recollection. Participants were also more likely to believe their recollections were accurate when they included knowledge of the setting and when the recollections provided a coherent story. Thus these experiments demonstrate that not only are factors such as vivid imagery important, but that the memory must be coherent for it to be believed—a fact that the SMS model highlights in particular, in terms of both consistency and fitting in with a person's life story.

A modified Metacognitive Model of autobiographical belief and memory

A modified version of the Metacognitive Model is now presented (Figure 15). The revised model incorporates elements of the SMS model, highlights the role of familiarity, and importantly, uses the findings of this thesis to emphasise the influential role of external evidence on belief and memory. All modifications are shown in red and are discussed below.

1. Retrieval, monitoring and long-term memory

The Metacognitive Model proposes that mental content is retrieved and monitored from long-term memory. In the modified version of the model presented in Figure 15, this process is further defined by incorporating aspects of the SMS model. Here, the retrieval process is controlled by the working-self. The working-self determines how wide or narrow the search for mental content is, and generates retrieval models to monitor the process and facilitate the source monitoring process. Thus it is important to show the retrieval and monitoring processes working together in unison. In the original model, retrieval and monitoring were only linked indirectly through long-term memory. In this modified version they are now directly linked and influenced by each other.

2. The role of other information and external evidence

The working-self also draws on other information, often from external sources (e.g., family members, photographs). In Experiments 1-4, information was retrieved about the cheating from the external fabricated evidence provided by the computer software. People are more likely to turn to external evidence when the long-term memory or autobiographical knowledge of the event is

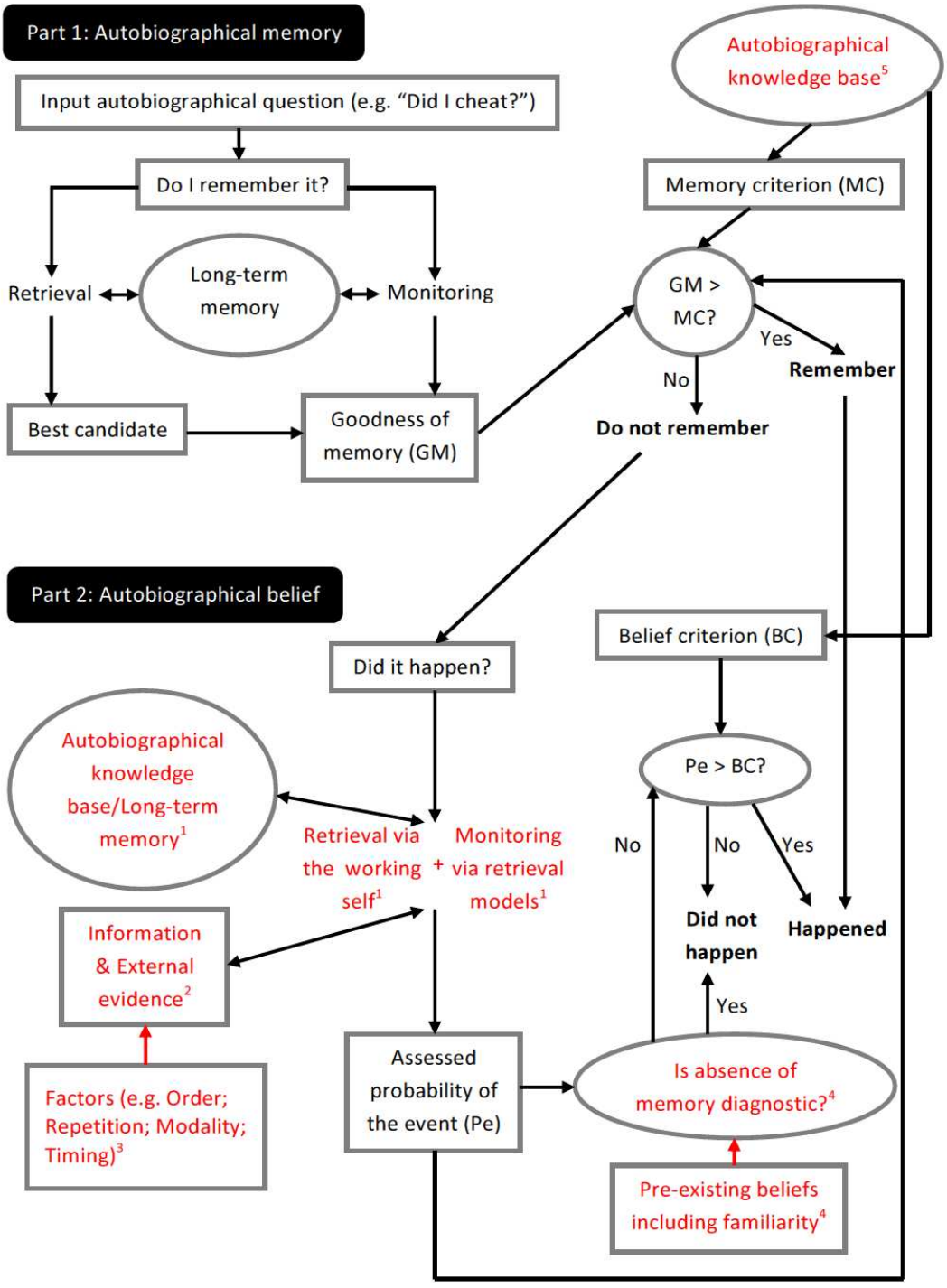

Figure 15. A modified version of the Metacognitive Model of autobiographical memory and belief.

unclear, than when the original memory remains strong and vivid. For instance, this can be seen in the late- vs. early-evidence conditions in Experiment 1, in which participants were more likely to turn to—and believe—false evidence presented after a delay than when it was presented immediately after the event. Thus the relative input and influence of information from long-term memory and external information is likely to vary from search to search.

It is important to note the reciprocal roles of long-term memory and external evidence with retrieval and monitoring processes. In Experiment 3 for instance, text likely facilitated a wide—yet focussed—search for material relevant to the cheating scenario. Whereas, photographs likely prompted a narrow—and thus more unsuccessful—search for material.

3. The role of mediating and moderating factors

This thesis demonstrates that the effect of external evidence is mediated and moderated by a series of factors. These factors affect how much emphasis is placed on the external evidence. For instance, Experiment 4 demonstrated that not all external evidence is considered. Rather, only the first piece of evidence appears to influence belief and memory construction.

4. Diagnosticity and the role of pre-existing beliefs

In Mazzoni and Kirsch's (2002) original version of the Metacognitive Model, a diagnosticity judgment about mental content (e.g., "Would I remember this, if it happened?") is made before any internal or external information is evaluated (see Figure 1 in Chapter 2). In this case, having no memory of cheating

usually indicates that the cheating did in fact not happen. Thus the false evidence is rejected and no more processing takes place. The evidence is only evaluated if a pre-existing belief—for instance, about childhood amnesia, distinctiveness of content, or memory degradation over time—is deemed relevant. This is consistent with the findings from Experiment 1, whereby the belief that memory fades over time presumably prompted participants to evaluate late-evidence, because having no memory simply meant they might have forgotten the event, rather than being diagnostic that the event did not occur.

However, the results of this thesis suggest two modifications to the diagnosticity judgment. First, findings from Experiments 2-4 suggest that there must have also been a pre-existing belief about familiarity, such as "familiar information is likely to be true". Thus, repeated or textual information was then further evaluated against the belief criterion. Although the original Metacognitive Model highlighted the role of pre-existing beliefs, familiarity itself as a mechanism was not specifically discussed or included in the diagram. Second, the diagnosticity judgment is likely to be influenced by—and thus made after, rather than before—both the evidence and probability of the event have been considered. For instance, in Experiment 3, the text evidence presumably made the mental image of cheating appear fluent and familiar, and thus the perceived probability of the event having occurred was increased. This probability assessment then likely influenced the diagnosticity judgment. It is important to note here that future research should focus on directly testing this mechanism.

5. The role of the autobiographical knowledge base in direct retrieval

Not all information is systematically processed. As the Memory Attribution Approach proposes, familiarity might be used as a heuristic. This means that in the stressful situation of being falsely accused, suspects might not engage effortful decision making processes.

Instead, they might use the feeling of familiarity to automatically assume that the event occurred, because a highly familiar mental event is likely to meet the belief or memory criteria. Thus all of the "Do I remember?" and "Did it happen?" stages might in fact be bypassed, and instead, the suspect either directly believes or "remembers" the event.

Indeed, people may rely on a range of heuristics other than familiarity to determine if an event is real or imagined. Usually, these heuristics provide a quick and efficient way of processing information, but if the two sources of information are highly similar this can lead to erroneous judgments. For instance, people may show a "bias towards the real" when deciding if something is real or imagined, based on the phenomenological qualities of the mental content (A. Kelly et al., 2002). Participants may reason, "If the item is vivid, it must be real" to judge vividly recalled and perceptually detailed content as real, whether or not it had actually been imagined.

This fits with what Conway refers to as *spontaneous recall* in the SMS model, as discussed in Chapter 2 (Conway & Pleydell-Pearce, 2000). Here, a highly specific cue—such as the presentation of false evidence—directly activates the autobiographical knowledge base. If the mental content meets heuristics such as "If the item is familiar or vivid, it must be real" then the criteria are automatically fulfilled, thus resulting in a false autobiographical belief and

perhaps even memory. Previously, the Metacognitive Model did not appear to emphasise this phenomenon. However, by adding this stage to the model, heuristic processing is able to account for not only spontaneous recall, but also a spontaneous belief that the event occurred. Although it is important to note that the experiments in this thesis do not directly test or show this mechanism, the results suggest that this is one way in which familiarity might influence belief and memory.

Chapter Summary

The findings from the four experiments comprising this thesis can be used to suggest modifications to current theoretical understanding of how false beliefs and memories are constructed. In the first part of this chapter, theoretical interpretations from the findings of this thesis were discussed. In the second part, Mazzoni and Kirsch's (2002) Metacognitive Model was modified to reflect these theoretical interpretations and incorporate aspects of other theories of memory. The modified model proposes four main changes. First, the model specifically shows how retrieval and monitoring processes are influenced by the working-self. Second, the model now suggests that a diagnosticity judgment is only made after the external evidence is considered, and that this judgment is influenced by familiarity. Third, the model depicts how people might come to spontaneously believe an event occurred or "relive" the event through direct-retrieval. Finally,—and most importantly in terms of this thesis—the new model includes the role of external evidence, and the mediating and moderating factors investigated in Experiments 1-4.

Chapter 10: General Discussion

"We believe that it is possible for scientific work to gain some knowledge about the reality of the world, by means of which we can increase our power and in accordance with which we can arrange our life."

—Sigmund Freud (1856-1939)

The previous chapter focussed on how the findings of this thesis suggest modifications to current autobiographical belief and memory theories. However, the findings are not limited to only having potential theoretical implications. This chapter also explores the practical and methodological implications of the experiments, as well as suggesting future research to refine the modified Metacognitive Model outlined in the previous chapter.

Part I: Practical implications

For each of the Experiments 1-4 in this thesis, the practical implications of the findings have been briefly discussed. The first part of this chapter looks at these implications—along with those from the digital editing questionnaire—in more depth and breadth, and poses some practical questions for future research. It should be noted that the findings from this thesis using the new driving task methodology provide suggestions and indications that should be used as a starting point for research, alongside other empirical research in the field. The

methodological advantages and limitations are discussed in more detail in the second part of this chapter.

Implications for legal processes

Throughout the experiments in this thesis, participants demonstrated compliance and false belief—and sometimes indicated that they had developed a false memory—when presented with misinformation and false evidence that they had committed a "crime". As shown in Experiments 1 and 3, the effects of false evidence were often over and above those of misinformation. These basic findings suggest that police should never present suspects with false evidence during interrogation, and that manuals (e.g., Inbau et al., 2005) should not encourage the use of this tactic. Although the use of fabricated evidence in the criminal justice system is rare, there are some cases in which it has been implicated. For instance, in 2001 in Sweden, a rioter was shot and arrested for having thrown stones at police officers. The rioter was then convicted using CCTV evidence. However, it was later revealed that the evidence had been altered to make the riot appear more aggressive and the officer involvement less brutal (see Maghan, O'reilly, & Shon, 2002). This thesis adds to the literature that shows that fabricated evidence is dangerous, and if used during interrogation, may lead to an increase in false confessions, and ultimately wrongful conviction.

More specifically, the results of the experiments in this thesis have implications for when evidence is disclosed during interrogation. For instance, Experiment 1 suggests that it is dangerous to present false evidence after a delay, because it can increase false belief and justification. Experiment 2 suggests that

this effect is exacerbated when the evidence is repeated. This is a worrying finding considering the fact that—as discussed in Chapter 4—there has been a shift towards presenting evidence late-on in an interrogation as opposed to early. This is thought to be a promising development, because information that the suspect provides throughout the interrogation will not have been contaminated by the evidence, and suggestible suspects should be less likely to confess (Moston & Engelberg, 1993). However, this is exactly the opposite of what this thesis shows and suggests that innocent suspects may be at risk when presented with misleading evidence late-on in interrogation.

Importantly, late disclosure of evidence is thought to induce diagnostic cues to deception (Granhag, 2010; Granhag, Strömwall, Willén, & Hartwig, 2012; Jordan, Hartwig, Wallace, Dawson, & Xhihani, 2012). In one study (Hartwig et al., 2006), students either did or did not commit a mock crime. Innocents were told to tell the truth and guilty students were told to lie and deny the crime. Police trainees—either trained or untrained in strategic evidence disclosure—then had to make veracity judgments after interviewing the students. Trained interviewers tended to disclose evidence late on in the interview and obtained an 85.4% accuracy rate compared to just 56.1% for the untrained interviewers. Another similar study involved undergraduate students making veracity judgments for videotaped mock-crime interrogations (Hartwig et al., 2005). Investigators either presented evidence against the student early or late in the interview. Liars in the early condition were detected at chance level, whereas liars in the late condition were detected above chance at 61.7%. The authors suggest that liars tend to have an increased cognitive demand. They omit information about evidence wherever

possible, and so appear more inconsistent than truth tellers when the evidence is presented later on. Interviewers and observers appear to use this inconsistency as a deception cue and can detect liars more accurately, although the same is not true for detecting truthful suspects. Other studies suggest that it is even more beneficial to disclose evidence incrementally throughout the interview (Dando & Bull, 2011; Sandham et al., 2011), and that this can be done using only one piece of evidence, but increasing its perceived strength throughout (Granhag et al., 2012). However, a recent study showed that late disclosure of evidence elicits more verbal cues to deception than gradual disclosure, and that gradual disclosure might even put innocents at risk of appearing guilty (Sorochinski et al., 2013).

Although these disclosure experiments are useful because they investigate the effects of both true and false evidence disclosure, they have tended to only focus on deception cues and veracity judgments, and thus the effects of the techniques on confession rates, beliefs and memories have remained largely unknown until this thesis. An exception to this is one study, in which there were found to be no differences between true and false confession rates for early and late disclosure of evidence (Jordan et al., 2012). However, the authors acknowledge that this result is likely because of a weak incentive to confess, and thus a low overall confession rate of 11%. This thesis suggests that—at least for innocent suspects—late disclosure of misleading evidence could be dangerous. Further research should also investigate the effects for guilty participants, and examine the mechanisms behind the effects of the phased disclosure approach with one or multiple pieces of evidence.

Indeed, timing of evidence presentation warrants further investigation because it appears to be such a complex factor with varying effects. Although Experiments 1 and 2 suggest that late-evidence disclosure might be dangerous, Experiment 3 suggests that this depends on its modality, and Experiment 4 suggests that early-evidence might be more powerful than late when multiple pieces are presented, and the most powerful modality is presented first. As discussed in Chapter 7, there appear to be interaction effects between repetition and modality in particular. These factors need to be investigated more fully for disclosure recommendations to be made. Perceived strength of evidence (e.g., Granhag et al., 2012) is also a factor that should be further investigated. Indeed, there is some research to suggest that evidence strength and police use of tactics are correlated. For instance, the amount of tactics police officers use during interrogation differs depending on how strong the evidence against the suspect is (Häkkänen, Ask, Kebbell, Alison, & Granhag, 2009). Indeed, when an interrogator expects a mock-suspect to be guilty, they use more interrogation techniques and exert more pressure to get a confession, yet ironically this is particularly the case with innocent suspects (Kassin, Goldstein, & Savitsky, 2003).

Although all of the experiments in this thesis emphasise the importance of timing of evidence presentation, Experiment 4 in particular showed the effects of confirmation bias, when people only accept information consistent with what they already know and believe, and reject any inconsistent information that they encounter afterwards. This experiment supports other research that shows that these cognitive biases are universal, and are not just made by particularly

vulnerable people. Even forensic experts and judges show confirmation bias when surrounded by influential evidence (see Kassin, Dror, & Kukucka, 2013 for a review). One study found that forensic examiners showed order effects even when using a randomized computer system, which suggests candidates for them to investigate. The experts spent less time examining potential candidates when they were lower down the list and were more likely to make false positives for candidates higher up the list (Dror, Wertheim, Fraser-Mackenzie, & Walajtys, 2012). Thus, people are capable of making a variety of cognitive biases when presented with evidence that although might be useful in most situations, can be highly dangerous in the context of the criminal justice system. The investigation of cognitive biases in context is becoming increasingly important, and is an idea that will be returned to later in the chapter.

Although the results of this thesis combined with the research discussed above appear to portray a somewhat negative view of legal implications in the criminal justice system, it must be emphasised that the use of false evidence is extremely rare. Also, there is a prevailing view that organisations such as the Police can initially be resistant to change because of their strong culture, yet any new implementations often become accepted and routine (Dror, Kassin, & Kukucka, 2013). This has been shown since PACE was introduced in the UK with mandatory recording of interviews, and also with recent eyewitness reforms (Garrett, 2013). Thus it is not impossible for psychologists, criminologists and sociologists to work together with Police to change current practice for the better.

Implications for advertising and the media

The findings from this thesis reveal some circumstances in which misleading advertisements might be particularly influential. For instance, Experiment 2 lends support to the idea that repetition is key when making people want to buy products. Contrary to previous research (e.g., Braun & Loftus, 1998), the findings of Experiment 3 suggest that misleading advertisements relying on text might be more influential than those with a more pictorial basis. The results of Experiment 4 further suggest that this might only be the case if misleading text is presented before misleading pictorial information. The interaction between modality of information and the order of its presentation is becoming increasingly important with the growing advertising market on the Internet, particularly through "banner ads" that are displayed around webpages. Interestingly, the modality of the webpage information itself might influence how much attention people pay to the surrounding advertisements. For instance, people are more likely to pay attention to banner ads when the webpage is video- or image-based as opposed to text- or image-and-text-based, and when the ads are shown on the first webpage they read as opposed to successive pages (Hsieh & Chen, 2011). Future research should investigate the interaction between the modality and order of advertising information with that of surrounding webpage information.

The findings from the digital editing questionnaire—discussed in Chapter 8—suggest that even if people are aware that advertisements or other digital materials in the media might have been altered, they are still likely to trust them. This potentially worrying finding is somewhat counterbalanced by the suggestion that educating people about digital editing might prompt them to consider digital alteration more in the future. One way of doing this might be to employ a symbol

or icon to notify readers of any digital alterations. Indeed, a "photation mark" could be a way of providing epistemic honesty to help readers evaluate the truthfulness of a digital image (Fosdick & Fahmy, 2007). The results from Chapter 8 also suggest that personalisation of digital materials might be key for educating participants about digital editing, because only participants who saw the doctored photographs themselves were likely to change their intentions towards digital media, rather than participants simply told about digital alteration. Thus experiments employing this technique are potentially highly useful educational tools. Future research should investigate people's assumptions before they are exposed to digital media and compare this to their views after exposure. In doing so, we could directly ascertain how people's assumptions change over time.

Part II: Methodological implications

This thesis has demonstrated that the new driving task methodology can be a highly useful tool for investigating false confessions, beliefs—and arguably—memories. Chapter 1 outlined a variety of false memory findings found using different methodologies in the field. Whilst these varied methods have been used to discover a wide range of interesting belief and memory phenomena, it was acknowledged in Chapter 3 that there was a need for a new and complementary method. The second part of this chapter discusses both the advantages and limitations of the new driving task method in the context of the current literature and findings from this thesis, and in doing so, highlights areas for future research.

Advantages of the driving task method

The primary reason that the new driving task methodology was valuable in this thesis, is because it was used to directly test the effects of a series of factors, some of which could not have otherwise been fully tested using other previous methods. For instance, no other current method is able to present participants with fabricated and personalised evidence immediately after the event, and thus no other method has been able to adequately test whether the timing of false evidence moderates or mediates its effect (Experiment 1). As discussed in Chapter 3, although some studies have used personalised doctored materials to influence participants' memories of recent events (e.g., Nash & Wade, 2009; Nash, Wade, & Lindsay, 2009), the experimenter had to laboriously doctor those materials and thus there was necessarily a time lag between sessions. The new method removed the lengthy doctoring process and only required a single experimental session, thus avoiding both attrition and contamination of participants' memories across sessions.

The evidence presented in Experiments 1-4 did not actually depict the participants themselves in the frame. Studies have shown that including details relevant to the self has a tendency to make false evidence more powerful (e.g., Desjardins & Scoboria, 2007; Hessen-Kayfitz & Scoboria, 2012; Strange et al., 2007). However, like other recent studies (Henkel, 2011a; Strange et al., 2007), Experiments 1-4 in this thesis showed that participants do not necessarily need to view themselves in the evidence to be persuaded by it. Simply having at least some similar details to the original event appears to be enough for false evidence to be influential. In terms of this thesis, all false evidence seen by participants was

at least partially true. The false evidence always used true details from the original event and simply changed the traffic light and/or clicking information. As Johnson (2006) suggests, it is always easier to induce a false memory when it is based on true information. Indeed, false memories are often formed from true memories to provide a coherent view of the self (Justice et al., 2012). The false evidence presented in the experiments always contained similar perceptual details that provided a "cognitive springboard" from which to imagine the cheating. These details presumably made the false evidence seem very similar to the original event, thus making it easy for the two to become confused. The details and features might also have "reactivated" similar perceptual details from the original event memory. This is known as the *feature importation process* (Lyle & Johnson, 2006; Weinstein & Nash, 2013) and explains why participants were likely confused when presented with both true and false details in a photograph in Experiment 3.

The driving task method also did not require confederates, who can sometimes be difficult to recruit and represent another possible confounding influence. For example, the memory implantation method often uses childhood photographs that must be provided by family members (e.g., Lindsay et al., 2004; Wade et al., 2002). The computer crash paradigm (e.g., Horselenberg et al., 2006; Kassin & Kiechel, 1996) also requires confederates to act as a witness and to probe participants in the waiting room. However, the new driving method allows for high levels of experimental control because of a complete lack of outside influence.

Initial concerns that participants might be highly suspicious and unconvinced by the set-up proved unfounded. Overall, only eight participants were excluded for expressing suspicion across all four experiments. Indeed, at debrief participants were often distinctly surprised and interested in the studies (e.g., "It really convinced me! After I saw the bunch of information telling me, I thought OK..."; "Interesting, very interesting! It completely deceived me! I believed the computer!"; "Oh! I was convinced!"). Thus, the new methodology is persuasive and convincing without being obviously "rigged".

The programming of the software means that it is relatively effortless to modify the procedure to investigate different factors and effects. This not only provides a coherent set of studies, but it also enables further work and potential comparison across studies. This is important, because as Mazzoni (2002) points out, different paradigms do not necessarily provide comparable results. Indeed, the DRM paradigm—as described in Chapter 1—uses only simple semantically related words and does not test autobiographical memory with a relevance of the "self". Although the findings from DRM studies can provide useful insights into memory—and indeed, are referred to in this thesis—they can be hard to compare with results obtained from other methods specifically studying autobiographical memory, such as the memory implantation method (e.g., Loftus & Pickrell, 1995). A recent study showed that although there is a small correlation between false memories produced using the DRM paradigm and those produced using the misinformation method (see Chapter 1), they generally rely on different mechanisms (Zhu, Chen, Loftus, Lin, & Dong, 2013).

Potential limitations of the driving task method

It is possible that the experimental control that the new method provides might have in itself diminished ecological validity. The results of this thesis are based on an accusation of cheating that meets experimental ethical requirements. Thus, the accusation differs substantially from complex acts involving intentionality (i.e. committing a crime) and from other real-world scenarios (i.e. police interrogation). As such, the results cannot be used to directly predict the likelihood of false confessions in the real world. However, as noted in Chapter 4, there were clear perceived consequences of the cheating and the experimenter was in a position of authority, thus making the scenario similar to a police interrogation situation. Also, it seems reasonable to assume that the mechanisms underlying the false evidence effects—and the factors involved—might contribute to varying compliant behaviours, illusory beliefs and memory distortions.

According to Leo (2008), laboratory studies—such as those demonstrated in this thesis—are useless in terms of studying false confessions. The results of experimental studies supposedly do not carry over into the real world because they are so far-removed. Indeed, Leo (2008) even claims that in real life, suspects might form false beliefs of committing crimes, but they do not form false memories like in experimental studies. However, there are several "real world" cases that dispute this claim. For instance, in the case of the Norfolk Four (see Chapter 1), one suspect developed false memories of gang-raping and murdering a young woman. In another criminal case, Paul Ingram developed vivid false memories of sexually abusing his children and conducting satanic rituals after repeated police interrogation. These cases suggest that experimental studies—

such as Experiments 1-4—might not in fact be so far-removed from real life phenomena, and instead might provide a useful tool for investigation. Indeed, a recent study has managed to bridge the gap between experiments and the real world by showing that people still make false claims and identifications and develop false beliefs and memories in a highly ecologically valid and stressful experimental setting (Morgan, Southwick, Steffian, Hazlett, & Loftus, 2013). In this study, US navy personnel were interrogated in a mock prisoner-of-war camp as part of their training. They were also deprived of both food and sleep, and exposed to propaganda. The propaganda videos the personnel were shown were varied as to who and what items were present at the time. At least 27% of personnel developed false memories of non-trivial items, such as weapons, being present. Some personnel were also exposed to a misleading photograph during the camp, and more than 50% of these personnel went on to then falsely identify their interrogator. Thus, this study shows that misinformation is highly persuasive, even when used in a highly stressful setting with people trained to withstand such influences.

Despite Leo's (2008) controversial realism concerns, his distinction between beliefs and memories is an important one, which has been emphasised throughout this thesis in terms of Mazzoni and Kirsch's (2002) Metacognitive Model. Arguably, the new driving task methodology does not directly test or demonstrate memory specifically. Rather, the procedure measures justification, which—as discussed earlier in Chapter 4—can sometimes overlap with confabulation. Studies have shown that confabulation over time can lead to the development of false memory (e.g., Drivdahl, Zaragoza, & Learned, 2009;

Zaragoza et al., 2001). Indeed, one study showed that reflecting on the meaning and implications of a suggested event—known as *conceptual elaboration*—often leads to a higher rate of false memory than elaborating on just perceptual details (Zaragoza, Mitchell, Payment, & Drivdahl, 2011). Presumably, elaboration links the suggested event to the corresponding true event, adding to the meaningfulness of the event and embedding the false memory in something which actually happened. Interestingly, the authors of this study suggest that false narratives might work in much the same way, by increasing semantic relations between the suggested event and the false event. This is yet another mechanism to be further investigated through which the false text in Experiments 3 and 4 of this thesis might have worked.

Recent research suggests that false memories might not only form from confabulation over time, but also that false memories themselves are persistent over time. Zhu et al. (2012) conducted a classic misinformation experiment (see Chapter 1) using a misleading narrative and followed participants up 1.5 years later. Half of the false memories persisted and the number of false memories originally developed was shown to correlate with the number found 1.5 years later. Importantly, false memory was maintained at the same rate as true memory. Thus this study suggests that—in line with the SMF as outlined in Chapter 2—once misattributed, a false memory has the same strength as a true memory and is likely to be constructed in much the same way. Indeed, Chan and McDermott (2006) even found that the experience of remembering is the same for false as for true memories. Thus, if participants in the experiments comprising this thesis did in fact develop a false memory of cheating from sometimes confabulating details

about the "crime", their false memory might persist despite debriefing and feel as real to them as a true memory (Clark et al., 2012). It would be interesting to follow-up on participants from Experiments 1-4 to see if this is in fact the case in several years time.

To complement the justification measure and provide a clearer indication of false memory, a remember/know/unsure was added to the new methodology in Experiments 3 and 4. Overall, this measure appeared to follow the same pattern of results as the other measures and indicated that a handful of participants across the two studies developed an actual false memory of the cheating. However, as mentioned in Chapter 7, the remember judgments might not actually be distinct from the know judgments. As Mazzoni and Kirsch (2002) suggest, both are indicative of at least believing that the person holds a memory for the event. Indeed, studies have shown that remember/know judgments tend to correlate more with belief than with recollection (e.g., Rubin et al., 2003; Rubin & Siegler, 2004). Chan and McDermott (2006) further posit that the judgments might not be "process-pure" because familiarity—which they equate with knowing—might influence judgments of recollection—which they equate with remembering—and so on. Thus remembering and knowing are very difficult to tease apart, and might be more informative when collapsed together (see Chapter 7). A measure of free recall might be a useful addition to the new methodology, which could then be coded separately.

Another potential limitation to this study was that true confessions were not investigated as well as false confessions. As described in Chapter 4, a "true cheater" manipulation was implemented several times during piloting but found to

be unsuccessful. Future research could potentially try other incentives to cheat using the current method that might prove more successful. Alternatively, the current driving task method could be used with a much larger sample size. Although several participants had to be excluded from each experiment because they did actually cheat, this sample was never large enough to conduct meaningful statistical analyses. However, a larger overall sample size might eliminate this problem. The descriptive data from Experiments 1-4 suggest that studying the behaviour of these "true cheaters" might be a useful addition to the current format. For example, 77% of true cheaters in Experiment 1 were judged to have fully believed they had cheated and 52% fully justified how it happened. These data are remarkably high and suggest that there is in fact a measurable difference between the cognitions and behaviours of cheaters and non-cheaters when accused. Real-life scenarios appear to support this premise, for instance with innocent suspects being more likely to waive their right to silence during interrogation than guilty suspects because of their faith in the justice system finding them innocent if they fully explain what happened (Kassin, 2005).

Indeed, other studies have used different methods to successfully induce both true and false confessions, and thus provide an interesting insight into overall underlying mechanisms. For example, in one study, participants were paired with a confederate but instructed to complete a pseudo-logic task alone (Russano et al., 2005). Some participants were left alone, whilst other participants were induced to cheat when the confederate approached them asking for help. All participants were accused of cheating by sharing answers with the confederate, and were asked to sign a confession statement. Some participants were interviewed by a

sympathetic investigator, who expressed concern, offered face-saving excuses, and suggested that it was in the participant's best interest to cooperate. Other participants were offered a deal, in which they were told by the investigator that if they agreed to sign the confession, the situation could be dealt with swiftly. However, if they did not agree to sign the statement, the professor would be called, likely resulting in more serious consequences. Although "guilty" participants were more likely to sign than "innocent" participants, both true and false confession rates increased when the investigator was either sympathetic or offered a deal relative to when no tactics were used. Thus the study demonstrated that widely used tactics were just as likely to make innocent suspects confess as guilty suspects. This finding is particularly worrying, given that mock-jurors are more likely to convict a suspect after a "sympathetic" interrogation (Kassin & McNall, 1991).

Another series of studies also manipulated whether or not participants cheated on a task and investigated how this affected participants' moral codes (Shu, Gino, & Bazerman, 2011). Interestingly, "guilty" participants displayed moral disengagement and tended to forget moral codes, whereas "innocent" participant recalled moral rules to a high level and displayed moral stringency. Indeed, when judging others' moral behaviours, one study found that memory for what happened becomes distorted dependent on the level of moral blame people attribute to the other person (Pizarro, Laney, Morris, & Loftus, 2006). Together, these studies suggest that investigating "true cheaters" using the new driving task methodology could provide a valuable insight into how guilty participants are affected by evidence presented in different ways.

It would also be useful to study individual differences using the new methodology. This was not the focus of this thesis and thus participants were not recruited or tested on this basis. Approximately equal numbers of males and females were tested across conditions in each of the four studies and age did not significantly vary. Thus, if there were any influential demographic variables, these should not have confounded effects of condition. Exploratory analyses indicated that these demographic variables only showed minimal correlations with dependent variables. Indeed, in all four experiments, males and females did not significantly differ in belief, and age and belief were not significantly correlated. This was apart from a Spearman's rho analysis that revealed a marginally significant negative correlation between age and belief in Experiment 4, $rs[120] = .16$, $p = .075$. However, squaring the correlation coefficients indicated an extremely weak correlation, such that only 3% of the variance was explained by this relationship. In Experiment 4 it was also found that males fully justified significantly more than females at 38% vs. 17% of the time, $\chi^2(1, N = 120) = 6.84$, $p = .009$, Cramer's V = .24. There was also a tendency for this to be the case in Experiment 1, $\chi^2(1, N = 75) = 3.57$, $p = .059$, Cramer's V = .22. Thus although minimal, there were findings that suggest that individual differences might correlate with false belief and influence justification.

Other studies more clearly demonstrate the influence of individual differences on false confession and belief rates, and show that there might in fact be interactions between different variables. For instance, when not stressed, women are more likely to falsely confess and develop beliefs than men, yet unlike men, they are no more likely to confess and believe when stressed (Forrest,

Wadkins, & Miller, 2002). Women are also more likely to falsely confess than men when the crime is highly plausible, but men's behaviour is more affected by personality variables such as compliance (Klaver, Lee, & Rose, 2008). Younger participants are more likely to confess than older participants (Redlich & Goodman, 2003), yet older adults are more likely than younger adults to erroneously recognize inferences when they have been repeated (McDermott & Chan, 2006). Indeed, older adults might be just as suggestible as juveniles, because they are more likely to have experienced a higher number of negative life events (Drake & Bull, 2011). Thus the current literature suggests that in terms of false confessions, demographics such as age and gender might be more influential than individual differences, such as personality or mood. Indeed, one study found no correlations between the likelihood of participants falsely confessing to a "crime" and factors such as suggestibility, fantasy-proneness and dissociation (Horselenberg et al., 2003). However, although they might not influence false confession rates, individual differences in depression, dissociation and personality might affect how accurate people themselves believe their own memories to be (Rubin et al., 2003; Rubin & Siegler, 2004).

The studies discussed above show that the relationships between individual difference variables themselves—and their influence on false confessions, beliefs and memories—are likely to be complex and interactive. The new driving task methodology could be used to test some of these relationships further, and perhaps after extensive testing it might be used to provide an experimental indication that a particular participant might be vulnerable during questioning. It could be used as a complementary measure alongside the

Gudjonsson Suggestibility Scale, which measures the extent that suspects yield to suggestive influence and how much they shift after interrogative pressure (Gudjonsson, 1984).

Part III: Theoretical considerations

As discussed in the previous chapter, the findings of this thesis suggest modifications to Mazzoni and Kirsch's (2002) Metacognitive Model of autobiographical belief and memory. Although the results extend current understanding of the role of familiarity in this model, the findings do not lend support to other mechanisms—such as credibility—as clearly. Previous research suggests that much of the false evidence effect is moderated by its perceived credibility (Nash, Wade, & Brewer, 2009). However, the results of Experiments 1-4 suggest that—at least when different factors of the false evidence are varied—credibility might act indirectly. For instance, in Experiment 1 it possible that the timing of the presentation of the evidence affected how credible and convincing the evidence was perceived to be. In Experiments 2 and 3, the increased familiarity brought about by a feeling of fluency could have influenced credibility. In Experiment 4, the consistency between the evidence and the original event memory might have influenced participants' perceptions of how credible the false evidence was. It is impossible to say whether credibility was influenced in these ways because the mechanism was not directly tested. Thus, future research should further investigate how credibility might operate indirectly, and how it might be influenced by other factors.

Other research also suggests that credibility might be an indirect mechanism behind the false evidence effect. For instance, nonprobative information—provided in either pictorial or verbal form—given alongside a written narrative, increases the narrative's perceived credibility (Newman et al., 2012). The authors suggest that the details provided in this information aid the generation of hypothesis-consistent material that then becomes accessible and fluent. In much the same way as in the experiments in this thesis, the feeling of fluency then likely influences perception of credibility. Indeed, the concept of consistency—and the generation of only hypothesis-consistent mental content—is now being recognised as an increasingly important mechanism behind the development of false memories. As Experiment 4 in this thesis demonstrated, people tend to show a confirmation bias such that they only accept information that is consistent with their initial "anchor".

The new driving task method could be further modified to explore the effects of consistency. For instance, what happens when participants are presented with one piece of evidence consistent with the accusation of cheating and one piece that is inconsistent and depicts no cheating whatsoever? According to the modified Metacognitive Model posited in the previous chapter, the credibility of the external evidence would be evaluated against other internal constructs, such as a belief criterion. When presented with consistent external evidence, the perceived credibility of this evidence might be higher than the internal belief that says "I didn't cheat", and thus participants should be more likely to falsely believe the evidence over their own memory. Consistent evidence would also prompt a broad search of memory for associated mental content, and thus likely result in

participants accepting the cheating accusation. In the inconsistent evidence condition however, the credibility of the cheating claim and the evidence might be decreased because of the conflict. In this case, the internal belief criterion saying "I didn't cheat" remains comparatively high to the credibility mechanism. Therefore in the inconsistent condition, participants might be less likely to falsely believe they cheated. The inconsistent evidence would also prompt only a narrow search of associated mental content, and thus likely result in rejection of the false accusation. This experiment would be theoretically interesting and would build on Experiment 4, in terms of whether or not consistency interacts with anchoring. Practically, the experiment would be useful because sometimes Police investigators might be faced with conflicting evidence—such as an alibi that conflicts with CCTV evidence—and must challenge the suspect and ask them to explain it. Also, the findings would complement current research examining what happens when jurors are presented with consistent vs. consistent evidence in court. This research suggests that there is a complex relationship between the type of evidence, experienced dissonance and resulting judgments (Ask, Reinhard, Marksteiner, & Granhag, 2011).

Chapter Summary

This thesis has used a novel experimental paradigm to investigate the false evidence effect and suggest modifications to current autobiographical memory theory. The findings have practical implications for police interrogation, advertising and the use of digitally altered images in the media. The development of the new driving task method also has methodological implications and raises

some interesting questions for future research. Finally, the findings highlight some important theoretical considerations, particularly for future work examining consistency of information.

Concluding remarks

This thesis has used science to suggest how and why people might develop false autobiographical beliefs and memories that can have both negative connotations and consequences. It is interesting now to take a much broader perspective and consider the issue of whether science should be used to "dampen" negative memories, whether they are true or false. Indeed, veracity itself does not seem to matter, because holding a false belief or memory that you have been traumatised produces the same emotional physiological stress responses (e.g., heart rate, skin conductance) and experiences as actual recollections of trauma (Chan & McDermott, 2006; McNally et al., 2004). Memory dampening might be beneficial for sufferers of Post-traumatic Stress Disorder, who can be plagued by "flashbacks" of a negative event for years afterwards. For instance, research suggests that a "cognitive vaccine"—in the form of a cognitively demanding visuospatial task completed shortly after the event—could reduce these traumatic intrusions (Holmes, James, Coode-Bate, & Deeprose, 2009). Yet interestingly, in a recent study when participants imagined experiencing a traumatic event, most said they would not take a drug to dampen the memory if offered one (Newman, Berkowitz, Nelson, Garry, & Loftus, 2011). It appears that memories, whether bad or good, form an intricate part of the self and identity as Conway suggests (e.g., Conway & Pleydell-Pearce, 2000; Conway et al., 2004). Perhaps there is an argument that people should at least be given the choice because everyone has an individual right to dampen private memories for self-preservation and good health. However, this could be pitted against a societal obligation not to dampen memory under certain circumstances, for instance if it means that a victim might

be able to identify the perpetrator of a crime (Kolber, 2006). These somewhat philosophical issues are likely to become more important and topical in the future with the advent of increased scientific understanding of human memory and improved trauma therapies.

References

ACPO. (2005). *Practice advice on core investigative doctrine*. Wyboston, UK: National Centre for Policing Excellence.

ACPO. (2009). *National Investigative Interview Strategy*. Wyboston, UK: National Police Improvement Agency

Alexander the Great. (n.d.). BrainyQuote.com. Retrieved August 22, 2013, from http://www.brainyquote.com/quotes/quotes/a/alexandert148611.html

Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review, 13*, 219-235. doi: 10.1177/1088868309341564

Appleby, S. C., Hasel, L. E., & Kassin, S. M. (2011). Police-induced confessions: an empirical analysis of their content and impact. *Psychology, Crime & Law iFirst*, 1-18. doi: 10.1080/1068316X.2011.613389

Ask, K., Reinhard, M. A., Marksteiner, T., & Granhag, P. A. (2011). Elasticity in evaluations of criminal evidence: Exploring the role of cognitive dissonance. *Legal and Criminological Psychology, 16*, 289-306. doi: 10.1348/135532510X510153

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge, UK: Cambridge University Press.

Basil, M. D. (1994). Multiple Resource Theory I: Application to Television Viewing. *Communication Research, 21,* 177-207. doi: 10.1177/009365094021002003

Begg, I. M., Anas, A., & Farinacci, S. (1992). Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General, 121,* 446-458. doi: 10.1037/0096-3445.121.4.446

Belli, R. F., Lindsay, D. S., Gales, M. S., & McCarthy, T. T. (1994). Memory impairment and source misattribution in postevent misinformation experiments with short retention intervals. *Memory & Cognition, 22*, 40-54. doi: 10.3758/BF03202760

Bernstein, D. M., Laney, C., Morris, E. K., & Loftus, E. F. (2005). False memories about food can lead to food avoidance. *Social Cognition, 23*, 11-34. doi: 10.1521/soco.23.1.11.59195

Bowers, J. M., & Bekerian, D. A. (1984). When will postevent information distort eyewitness testimony? *Journal of Applied Psychology, 69*, 466-472. doi: 10.1037/0021-9010.69.3.466

Braun, K. A., Ellis, R., & Loftus, E. F. (2002). Make my memory: How advertising can change our memories of the past. *Psychology & Marketing, 19*, 1-23. doi: 10.1002/mar.1000

Braun, K. A., & Loftus, E. F. (1998). Advertising's misinformation effect. *Applied Cognitive Psychology, 12*, 569-591. doi: 10.1002/(SICI)1099-0720(1998120)12:6<569::AID-ACP539>3.0.CO;2-E

Braun-LaTour, K. A., LaTour, M. S., Pickrell, J. E., & Loftus, E. F. (2004). How and when advertising can influence memory for consumer experience. *Journal of Advertising, 33,* 7-25.

Bremner, J. D., Shobe, K. K., & Kihlstrom, J. F. (2000). False memories in women with self-reported childhood sexual abuse: An empirical study. *Psychological Science, 11*, 333-337. doi: 10.1111/1467-9280.00266

Brewer, W. F. (1977). Memory for the pragmatic implications of sentences. *Memory & Cognition, 5*, 673-678. doi: 10.3758/BF03197414

Bryant, R. (Ed.). (2006). *Blackstone's student police officer handbook*. Oxford, England: Oxford University Press.

Burt, C., Kemp, S., & Conway, M. A. (2004). Memory for true and false autobiographical event descriptions. *Memory, 12*, 545-542. doi: 10.1080/09658210344000071

Ceci, S. J., Loftus, E. F., Leichtman, M. D., & Bruck, M. (1994). The possible role of source misattribution in the creation of false beliefs among preschoolers. *International Journal of Clinical and Experimental Hypnosis, 42*, 304-320. doi: 10.1080/00207149408409361

Chan, J. C. K., & McDermott, K. B. (2006). Remembering pragmatic inferences. *Applied Cognitive Psychology, 20*, 633-639. doi: 10.1002/acp.1215

Chapman, G. B., & Bornstein, B. H. (1996). The more you ask for, the more you get: Anchoring in personal injury verdicts. *Applied Cognitive Psychology, 10*, 519-540. doi: 10.1002/(SICI)1099-0720(199612)10:6<519::AID-ACP417>3.0.CO;2-5

Chojnacki, D. E., Cicchini, M. D., & White, L. T. (2008). An empirical basis for the admission of expert testimony on false confessions. *Arizona State Law Journal, 40*, 1-45.

Christiaansen, R. E., & Ochalek, K. (1983). Editing misleading information from memory: Evidence for the coexistence of original and postevent information. *Memory & Cognition, 11*, 467-475. doi: 10.3758/BF03196983

Christie, A. (2001). *The Mysterious Affair at Styles*. London, UK: HarperCollins Publishers.

Chrobak, Q. M., & Zaragoza, M. S. (2008). Inventing stories: Forcing witnesses to fabricate entire fictitious events leads to freely reported false memories. *Psychonomic Bulletin & Review, 15,* 1190-1195. doi: 10.3758/PBR.15.6.1190

Clancy, S. A., McNally, R. J., Schacter, D. L., Lenzenweger, M. F., & Pitman, R. K. (2002). Memory distortion in people reporting abduction by aliens. *Journal of Abnormal Psychology, 111,* 455–461. doi: 10.1037//0021-843X.111.3.455

Clark, A., Nash, R. A., Fincham, G., & Mazzoni, G. A. L. (2012). Creating non-believed memories for recent autobiographical events. *PLoS ONE, 7*, e32998. doi: 10.1371/journal.pone.0032998

Clifasefi, S. L., Bernstein, D. M., Mantonakis, A., & Loftus, E. F. (2013). "Queasy does it": False alcohol beliefs and memories may lead to diminished alcohol preferences. *Acta Psychologica, 143*, 14-19. doi: 10.1016/j.actpsy.2013.01.017

Conway, M. A. (2005). Memory and the self. *Journal of Memory and Language, 53*, 594–628. doi: 10.1016/j.jml.2005.08.005

Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the Self-Memory System. *Psychological Review, 107,* 261-288. doi: 10.1037/0033-295X.107.2.261

Conway, M. A., Singer, J. A., & Tagini, A. (2004). The self and autobiographical memory: Correspondence and coherence. *Social Cognition, 22*, 491-529. doi: 10.1521/soco.22.5.491.50768

Cornoldi, C. (1998). The impact of metacognitive reflection on cognitive control. In G. A. L. Mazzoni & T. Nelson (Eds.), *Metacognition and cognitive neuropsychology: Monitoring and control processes* (pp. 139-159). Hillsdale, NJ: Lawrence Erlbaum Associates

Crombag, H. F. M., Wagenaar, W. A., & van Koppen, P. J. (1996). Crashing memories and the problem of 'source monitoring'. *Applied Cognitive Psychology, 10,* 95-104. doi: 10.1002/(SICI)1099-0720(199604)10:2<95::AID-ACP366>3.0.CO;2-#

Cushman, F., Young, L., & Hauser, M. (2006). The role of conscious reasoning and intuition in moral judgment: Testing three principles of harm. *Psychological Science, 17,* 1082-1089. doi: 10.1111/j.1467-9280.2006.01834.x

Dando, C. J., & Bull, R. (2011). Maximising opportunities to detect verbal deception: Training police officers to interview tactically. *Journal of Investigative Psychology and Offender Profiling, 8*, 189–202. doi: 10.1002/jip.145

Desjardins, T., & Scoboria, A. (2007). "You and your best friend Suzy put Slime in Ms. Smollett's desk": Producing false memories with self- relevant

details. *Psychonomic Bulletin & Review, 14,* 1090-1095. doi: 10.3758/BF03193096

Desmarais, S. L., & Read, J. D. (2011). After 30 years, what do we know about what jurors know? A meta-analytic review of lay knowledge regarding eyewitness factors. *Law and Human Behavior, 35*, 200–210. doi: 10.1007/s10979-010-9232-6

Donnelly, J. (2004). *A Northern Light.* Orlando, FL: Harcourt Brace International.

Drake, K., & Bull, R. (2011). Individual differences in interrogative suggestibility: life adversity and field dependence *Psychology, Crime & Law, 17*, 677-687. doi: 10.1080/10683160903511967

Drivdahl, S. B., Zaragoza, M. S., & Learned, D. M. (2009). The role of emotional elaboration in the creation of false memories. *Applied Cognitive Psychology, 23,* 13-35. doi: 10.1002/acp.1446

Drizin, S. A., & Leo, R. A. (2004). The problem of false confessions in the post-DNA world *North Carolina Law Review, 82*, 891–1007

Dror, I. E. (2012). Cognitive bias in forensic science. *Science & Technology 2012 Yearbook,* 43-45.

Dror, I. E., Charlton, D., & Peron, A. (2006). Contextual information renders experts vulnerable to making erroneous identifications. *Forensic Science International, 156,* 74-78. doi: 10.1016/j.forsciint.2005.10.017

Dror, I. E., Kassin, S. M., & Kukucka, J. (2013). New application of psychology to law: Improving forensic evidence and expert witness contributions. *Journal of Applied Research in Memory and Cognition.*

Dror, I. E., Wertheim, K., Fraser-Mackenzie, P., & Walajtys, J. (2012). The impact of human-technology cooperation and distributed cognition in forensic science: Biasing effects of AFIS contextual information on human experts. *Journal of Forensic Sciences, 57*, 343-352. doi: 10.1111/j.1556-4029.2011.02013.x

Echterhoff, G., Hirst, W., & Hussy, W. (2005). How eyewitnesses resist misinformation: Social postwarnings and the monitoring of memory characteristics. *Memory & Cognition, 33*, 770-782. doi: 10.3758/BF03193073

Elaad, E., Ginton, A., & Ben-Shakhar, G. (1994). The effects of prior expectations and outcome knowledge on polygraph examiners' decisions. *Journal of Behavioral Decision Making, 7*, 279-292. doi: 10.1002/bdm.3960070405

Englich, B., Mussweiler, T., & Strack, F. (2006). Playing dice With criminal sentences: The influence of irrelevant anchors on experts' judicial decision making. *Personality and Social Psychology Bulletin, 32*, 188-200. doi: 10.1177/0146167205282152

Epley, N., & Gilovich, T. (2006). The anchoring-and-adjustment heuristic: Why the adjustments are insufficient. *Psychological Science, 17*, 311-318. doi: 10.1111/j.1467-9280.2006.01704.x

Festinger, L. (1957). *Conflict, decision, and dissonance*. Stanford, CA: Stanford University Press.

Field, A. P., Argyris, N. G., & Knowles, K. A. (2001). Who's afraid of the big bad wolf: A prospective paradigm to test Rachman's indirect pathways in

children *Behaviour Research and Therapy, 39*, 1259–1276 doi: 10.1016/S0005-7967(00)00080-2

Field, A. P., & Lawson, J. (2003). Fear information and the development of fears during childhood: Effects on implicit fear responses and behavioural avoidance. *Behaviour Research and Therapy, 41*, 1277-1293. doi: 10.1016/S0005-7967(03)00034-2

Forrest, K. D., Wadkins, T. A., & Miller, R. L. (2002). The role of preexisting stress on false confessions: An empirical study. *Journal of Credibility Assessment & Witness Psychology, 3*, 23-45.

Fosdick, S., & Fahmy, S. (2007). Epistemic honesty and the default assumption that photos are true. *SIMILE: Studies In Media & Information Literacy Education, 7*, 1-10. doi: 10.3138/sim.7.1.003

Foster, J. L., & Garry, M. (2012). Building false memories without suggestion. *The American Journal of Psychology, 125*, 225-232. doi: 10.5406/amerjpsyc.125.2.0225

Foster, J. L., Huthwaite, T., Yesberg, J. A., Garry, M., & Loftus, E. F. (2012). Repetition, not number of sources, increases both susceptibility to misinformationand confidence in the accuracy of eyewitnesses *Acta Psychologica, 139*, 320–326 doi: 10.1016/j.actpsy.2011.12.004

French, L., Garry, M., & Mori, K. (2011). Relative—not absolute—judgments of credibility affect susceptibility to misinformation conveyed during discussion. *Acta Psychologica, 136*, 119–128. doi: 10.1016/j.actpsy.2010.10.009

Frenda, S. J., Knowles, E. D., Saletanc, W., & Loftus, E. F. (2013). False memories of fabricated political events. *Journal of Experimental Social Psychology, 49*, 280-286. doi: 10.1016/j.jesp.2012.10.013

Freud, S. (2012). *The Future of an Illusion* (T. Dufresne Ed.). Peterborough, CA: Broadview Press Ltd.

Frost, P. (2000). The quality of false memory over time: Is memory for misinformation "remembered" or "known"? *Psychonomic Bulletin & Review, 7,* 531-536. doi: 10.3758/BF03214367

Gabbert, F., Memon, A., & Allan, K. (2003). Memory conformity: Can eyewitnesses influence each other's memories for an event? *Applied Cognitive Psychology, 17*, 533-543. doi: 10.1002/acp.885

Gallo, D. A. (2010). False memories and fantastic beliefs: 15 years of the DRM illusion. *Memory & Cognition 28*, 833-848. doi: 10.3758/MC.38.7.833

Garrett, B. L. (2008). Judging innocence. *Columbia Law Review, 108*, 55-142.

Garrett, B. L. (2010). The substance of false confessions. *Stanford Law Review, 62*, 1051-1119.

Garrett, B. L. (2013). Trial and error. In C. R. Huff & M. Killias (Eds.), *Wrongful convictions and miscarriages of justice: Causes and remedies in North American and European Criminal Justice Systems* (pp. 77). London, UK: Routledge.

Garry, M., French, L., Kinzett, T., & Mori, K. (2008). Eyewitness memory following discussion: Using the MORI technique with a Western sample. *Applied Cognitive Psychology, 22*, 431-439. doi: 10.1002/acp.1376

Garry, M., & Gerrie, M. P. (2005). When photographs create false memories. *Current Directions in Psychological Science, 14*, 321-325. doi: 10.1111/j.0963-7214.2005.00390.x

Garry, M., & Wade, K. A. (2005). Actually, a picture is worth less than 45 words: Narratives produce more false memories than photographs. *Psychonomic Bulletin & Review, 12*(2), 359-366.

Geraerts, E., Bernstein, D. M., Merckelbach, H., Linders, C., Raymaekers, L., & Loftus, E. F. (2008). Lasting false beliefs and their behavioral consequences. *Psychological Science, 19*, 749-753.

Gobbo, C. (2000). Assessing the effects of misinformation on children's recall: how and when makes a difference. *Applied Cognitive Psychology, 14*, 163-182. doi: 10.1002/(SICI)1099-0720(200003/04)14:2<163::AID-ACP630>3.0.CO;2-H

Goff, L. M., & Roediger, H. L. (1998). Imagination inflation for action events: Repeated imaginings lead to illusory recollections. *Memory & Cognition, 26*, 20-33. doi: 10.3758/BF03211367

Golda Meir. (n.d.). BrainyQuote.com. Retrieved August 22, 2013, from http://www.brainyquote.com/quotes/quotes/g/goldameir131891.html

Goldsmith, M., & Koriat, A. (1999). The strategic regulation of memory reporting: Mechanisms and performance consequences. In D. Gopher & A. Koriat (Eds.), *Attention and Performance XVII. Cognitive regulation of performance: Interaction of theory and application* (pp. 373-400). MA, Cambridge: MIT Press.

Goodman, G. S., Ogle, C. M., Block, S. D., Harrisa, L. S., Larson, R. P., Augusti, E. M., . . . Urquiza, A. (2011). False memory for trauma-related Deese–Roediger–McDermott lists in adolescents and adults with histories of child sexual abuse *Development and Psychopathology, 23,* 423-438. doi: 10.1017/S0954579411000150

Granhag, P. A. (2010). *Interviewing to elicit cues to detect deception.* Paper presented at the New ways of gathering information from witness and suspects: Proceedings of the Forensic Psychology Conference, University of Leicester, Leicester, UK.

Granhag, P. A., Strömwall, L. A., Willén, R. M., & Hartwig, M. (2012). Eliciting cues to deception by tactical disclosure of evidence: The first test of the Evidence Framing Matrix. *Legal and Criminological Psychology,* 1-15. doi: 10.1111/j.2044-8333.2012.02047.x

Greer, J. D., & Gosen, J. D. (2002). How much is too much? *Visual Communication Quarterly, 9,* 4-13. doi: 10.1080/15551390209363485

Gudjonsson, G. H. (1984). A new scale of interrogative suggestibility. *Personality and Individual Differences, 5,* 303-314. doi: 10.1016/0191-8869(84)90069-2

Gudjonsson, G. H. (2003). *The Psychology of Interrogations and Confessions: A Handbook.* Chichester, UK: Wiley.

Hartmann, M. (2009, July 8). Kelly Clarkson Slimmed Down On *Self* Via Photoshop. Retrieved August 22, 2013, from http://jezebel.com/5332409/kelly-clarkson-slimmed-down-on-self-via-photoshop

Hartwig, M., Granhag, P. A., Strömwall, L. A., & Kronkvist, O. (2006). Strategic use of evidence during police interviews: When training to detect deception works. *Law and Human Behavior, 30,* 603-619. doi: 10.1007/s10979-006-9053-9

Hartwig, M., Granhag, P. A., Strömwall, L. A., & Vrij, A. (2005). Detecting Deception Via Strategic Disclosure of Evidence. *Law and Human Behavior, 29,* 469-484. doi: 10.1007/s10979-005-5521-x

Hasel, L. E., & Kassin, S. M. (2009). On the presumption of evidentiary independence : Can confessions corrupt eyewitness identifications? *Psychological Science, 20,* 122-126. doi: 10.1111/j.1467-9280.2008.02262.x

Hasher, L., Goldstein, D., & Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behavior, 16,* 107–112. doi: 10.1016/S0022-5371(77)80012-1

Hawkins, S. A., & Hoch, S. J. (1992). Low-involvement learning: Memory without evaluation. *Journal of Consumer Research, 19,* 212-225. doi: 10.1086/209297

Henkel, L. A. (2008). Jurors' reactions to recanted confessions: Do the defendant's personal and dispositional characteristics play a role? *Psychology, Crime & Law, 14,* 565-578. doi: 10.1080/10683160801995247

Henkel, L. A. (2011a). Photograph-induced memory errors: When photographs make people claim they have done things they have not. *Applied Cognitive Psychology, 25,* 78-86. doi: 10.1002/acp.1644

Henkel, L. A. (2011b). Seeing photos makes us read between the lines: The influence of photos on memory for inferences. *The Quarterly Journal of Experimental Psychology*, *65*, 773-795. doi: 10.1080/17470218.2011.628400

Henkel, L. A., & Carbuto, M. (2008). How source misattributions arise from verbalization, mental imagery, and pictures. In M. Kelley (Ed.), *Applied Memory* (pp. 213-234). Hauppauge, NY: Nova Science.

Henkel, L. A., Coffman, K. A. J., & Dailey, E. M. (2008). A survey of people's attitudes and beliefs about false confessions. *Behavioral Sciences & the Law, 26*, 555-584. doi: 10.1002/bsl.826

Henkel, L. A., & Coffman, K. J. (2004). Memory distortions in coerced false confessions: A source monitoring framework analysis. *Applied Cognitive Psychology, 18*, 567-588. doi: 10.1002/acp.1026

Hessen-Kayfitz, J. K., & Scoboria, A. (2012). False memory is in the details: Photographic details differentially predict memory formation. *Applied Cognitive Psychology, 26*, 333-341. doi: 10.1002/acp.1839

Holmes, E. A., James, E. L., Coode-Bate, T., & Deeprose, C. (2009). Can playing the computer game "Tetris" reduce the build-up of flashbacks for trauma? A Proposal fromCognitive Science. *PLoS ONE, 4*, e413. doi: 10.1371/journal.pone.0004153

Horry, R., Palmer, M. A., Sexton, M. L., & Brewer, N. (2012). Memory conformity for confidently recognized items: The power of social influence on memory reports. *Journal of Experimental Social Psychology, 48*, 783–786. doi: 10.1016/j.jesp.2011.12.010

Horselenberg, R., Merckelbach, H., & Josephs, S. (2003). Individual differences and false confessions: A conceptual replication of Kassin and Kiechel (1996). *Psychology, Crime & Law, 9,* 1-8. doi: 10.1080/10683160308141

Horselenberg, R., Merckelbach, H., Smeets, T., Franssens, D., Peters, G.-J. Y., & Zeles, G. (2006). False confessions in the lab: Do plausibility and consequences matter? *Psychology, Crime & Law, 12,* 61-75. doi: 10.1080/1068310042000303076

Hsieh, Y. C., & Chen, K. H. (2011). How different information types affect viewer's attention on internet advertising. *Computers in Human Behavior, 27,* 935-945. doi: 10.1016/j.chb.2010.11.019

http://www.innocenceproject.org

Huang, E. S. (2001). Readers' perception of digital alteration in photojournalism. *Journalism & Communication Monographs, 3,* 147-182. doi: 10.1177/152263790100300302

Hyman, I. E., & Billings, F. J. (1998). Individual differences and the creation of false childhood memories. *Memory, 6,* 1-20. doi: 10.1080/741941598

Hyman, I. E., & Kleinknecht, E. E. (1999). False childhood memories: Research, theory, and applications. In L. M. Williams & V. L. Banyard (Eds.), *Trauma and memory* (pp. 178-188). Thousand Oaks, CA: Sage.

Häkkänen, H., Ask, K., Kebbell, M., Alison, L., & Granhag, P. A. (2009). Police officers' views of effective interview tactics with suspects: The effects of weight of case evidence and discomfort with ambiguity. *Applied Cognitive Psychology, 23,* 468–481. doi: 10.1002/acp.1491

Inbau, F. E., Reid, J., & Buckley, J. (1986). *Criminal interrogation and confessions* (3rd ed.). Baltimore, MD: Williams & Wilkins.

Inbau, F. E., Reid, J. E., Buckley, J. P., & Jayne, B. C. (2005). *Essentials of the Reid technique: Criminal interrogation and confessions*. Sudbury, MA: Jones and Bartlett Publishers.

Irving, B. L., & Hilgendorf, L. (1980). *Police interrogation: The psychological approach*. Royal Commission on Criminal Procedure, Research Study No. 1. London: HMSO.

Jacoby, L. L., & Kelley, C. M. (1987). Unconscious influences of memory for a prior event. *Personality and Social Psychology Bulletin, 13,* 314-336. doi: 10.1177/0146167287133003

Jacoby, L. L., Kelley, C. M., & Dywan, J. (1989). Memory attributions. In H. L. I. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 391-422). Hillsdale, NJ: Lawrence Erlbaum Associates.

Jacoby, L. L., Woloshyn, V., & Kelley, C. (1989). Becoming famous without being recognized: Unconscious influences of memory produced by dividing attention. *Journal of Experimental Psychology: General, 118,* 115-125. doi: 10.1037/0096-3445.118.2.115

Jamison, J., Karlan, D., & Schechter, L. (2008). To deceive or not to deceive: The effect of deception on behavior in future laboratory experiments. *Journal of Economic Behavior & Organization, 68,* 477–488. doi: 10.1016/j.jebo.2008.09.002

Jenkinson, P. M., & Conway, M. A. (2012). Memory and the self. In A. Fotopoulou, D. Pfaff & M. A. Conway (Eds.), *From the Couch to the Lab: Trends in Psychodynamic Neuroscience* (pp. 386-401). Oxford, UK: Oxford University Press.

Johnson, M. K. (2006). Memory and reality. *American Psychologist, 61,* 760-771. doi: 10.1037/0003-066X.61.8.760

Johnson, M. K., Foley, M. A., Suengas, A. G., & Raye, C. L. (1988). Phenomenal characteristics of memories for perceived and imagined autobiographical events. *Journal of Experimental Psychology: General, 117,* 371-376. doi: 10.1037/0096-3445.117.4.371

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114*, 3-28. doi: 10.1037/0033-2909.114.1.3

Johnson, M. K., Kahan, T. L., & Raye, C. L. (1984). Dreams and reality monitoring. *Journal of Experimental Psychology: General, 113,* 329-344. doi: 10.1037/0096-3445.113.3.329

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88*, 67-85. doi: 10.1037/0033-295X.88.1.67

Jones, T. C., Jacoby, L. L., & Gellis, L. A. (2001). Cross-modal feature and conjunction errors in recognition memory. *Journal of Memory and Language, 44*, 131-152. doi: 10.1006/jmla.2001.2713

Jordan, S., Hartwig, M., Wallace, B., Dawson, E., & Xhihani, A. (2012). Early versus late disclosure of evidence: Effects on verbal cues to deception, confessions, and lie catchers' accuracy. *Journal of Investigative Psychology and Offender Profiling, 9*, 1-2. doi: 10.1002/jip.1350

Justice, L. V., Morrison, C. M., & Conway, M. A. (2012). True and intentionally fabricated memories. *The Quarterly Journal of Experimental Psychology, iFirst*, 1-8. doi: 10.1080/17470218.2012.734832

Kalbfleisch, P. J. (1994). The language of detecting deceit. *Journal of Language and Social Psychology, 13*, 469-496. doi: 10.1177/0261927X94134006

Kassin, S. M. (1997). False memories turned against the self. *Psychological Inquiry, 8*, 300-302. doi: 10.1207/s15327965pli0804_4

Kassin, S. M. (2005). On the psychology of confessions: Does innocence put innocents at risk? *American Psychologist, 60*, 215-228. doi: 10.1037/0003-066X.60.3.215

Kassin, S. M., Bogart, D., & Kerner, J. (2012). Confessions that corrupt: Evidence from the DNA exoneration case files. *Psychological Science, 23*, 41-45. doi: 10.1177/0956797611422918

Kassin, S. M., Dror, I. E., & Kukucka, J. (2013). The forensic confirmation bias: Problems, perspectives, and proposed solutions. *Journal of Applied Research in Memory and Cognition*.

Kassin, S. M., Goldstein, C. C., & Savitsky, K. (2003). Behavioral confirmation in the interrogation room: On the dangers of presuming guilt. *Law and Human Behavior, 27*, 187-203. doi: 10.1023/A:1022599230598

Kassin, S. M., & Kiechel, K. L. (1996). The social psychology of false confessions: Compliance, internalization, and confabulation. *Psychological Science, 7*, 125-128. doi: 10.1111/j.1467-9280.1996.tb00344.x

Kassin, S. M., Leo, R. A., Meissner, C. A., Richman, K. D., Colwell, L. H., Leach, A., & La Fon, D. (2007). Police interviewing and interrogation: A self-report survey of police practices and beliefs. *Law and Human Behavior, 31,* 381–400. doi: 0.1007/s10979-006-9073-5

Kassin, S. M., & McNall, K. (1991). Police interrogations and confessions: Communicating promises and threats by pragmatic implication. *Law and Human Behavior, 15*, 233-251. doi: 10.1007/BF01061711

Kassin, S. M., & Sukel, H. (1997). Coerced confessions and the jury: An experimental test of the "harmless error" rule. *Law and Human Behavior, 21*, 27-46. doi: 10.1023/A:1024814009769

Kassin, S. M., & Wrightsman, L. S. (1985). Confession evidence. In S. M. Kassin & L. S. Wrightsman (Eds.), *The psychology of evidence and trial procedure* (pp. 67-94). Beverly Hills, CA: Sage Publications.

Kelly, A., Carroll, M., & Mazzoni, G. A. L. (2002). Metamemory and reality monitoring. *Applied Cognitive Psychology, 16*, 407-428. doi: 10.1002/acp.803

Kelly, J. E., & Nace, D. (1994). Digital imaging & believing photos. *Visual Communication Quarterly, 1*, 4-18. doi: 10.1080/15551393.1994.10387481

Kennedy, G. E., Judd, T. S., Churchward, A., Gray, K., & Krause, K. (2008). First year students' experiences with technology: Are they really digital natives? *Australasian Journal of Educational Technology, 24*, 108-122.

Klaver, J. R., Lee, Z., & Rose, V. G. (2008). Effects of personality, interrogation techniques and plausibility in an experimental false confession paradigm.

Legal and Criminological Psychology, 13, 71–88. doi: 10.1348/135532507X193051

Kolber, A. J. (2006). Therapeutic forgetting: The legal and ethical implications of memory dampening *Vanderbilt Law Review, 59,* 1561-1626.

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review, 103,* 490-517. doi: 10.1037/0033-295X.103.3.490

Krupat, E., & Garonzik, R. (1994). Participants' expectations and the search for alternatives to deception in social psychology. *The British Journal of Social Psychology, 33,* 211-222. doi: 10.1111/j.2044-8309.1994.tb01019.x

Lane, S. M., & Zaragoza, M. S. (1995). The recollective experience of cross-modality confusion errors. *Memory & Cognition, 23,* 607–610. doi: 10.3758/BF03197262

Laney, C., Fowler, N. B., Nelson, K. J., Bernstein, D. M., & Loftus, E. F. (2008). The persistence of false beliefs. *Acta Psychologica, 129,* 190-197. doi: 10.1016/j.actpsy.2008.05.010

Laney, C., Morris, E. K., Bernstein, D. M., Wakefield, B. M., & Loftus, E. F. (2008). Asparagus, a love story: Healthier eating could be just a false memory away. *Experimental Psychology, 55,* 291-300. doi: 10.1027/1618-3169.55.5.291

Leo, R. A. (1996). Inside the interrogation room. *Journal of Criminal Law and Criminology, 86,* 266-303.

Leo, R. A. (2008). *Police interrogation and American justice*. Cambridge, MA: Harvard University Press.

Leo, R. A., & Liu, B. (2009). What do potential jurors know about police interrogation techniques and false confessions? *Behavioral Sciences & the Law, 27*, 381-399. doi: 10.1002/bsl.872

Leo, R. A., & Ofshe, R. (1998). The consequences of false confessions: Deprivations of liberty and miscarriages of justice in the age of psychological interrogation *Journal of Criminal Law and Criminology, 88,* 429-496. doi: 10.2307/1144288

Leonardo da Vinci. (n.d.). BrainyQuote.com. Retrieved August 22, 2013, from http://www.brainyquote.com/quotes/quotes/l/leonardoda140595.html

Lindsay, D. S. (1990). Misleading suggestions can impair eyewitnesses' ability to remember event details. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 1077-1083. doi: 10.1037/0278-7393.16.6.1077

Lindsay, D. S. (2008). Source monitoring. In H. L. Roediger (Ed.), *Cognitive psychology of memory* (Vol. 2, pp. 325-348). Oxford, England: Elsevier.

Lindsay, D. S., Hagen, L., Read, J. D., Wade, K. A., & Garry, M. (2004). True photographs and false memories. *Psychological Science, 15*, 149-154. doi: 10.1111/j.0956-7976.2004.01503002.x

Loftus, E. F. (1977). Shifting human color memory. *Memory & Cognition, 5*, 696-699. doi: 10.3758/BF03197418

Loftus, E. F. (1979). *Eyewitness testimony*. Cambridge, MA: Harvard University Press.

Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory, 12*, 361-366. doi: 10.1101/lm.94705

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory, 4*, 19-31. doi: 10.1037/0278-7393.4.1.19

Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior, 13*, 585-589. doi: 10.1016/S0022-5371(74)80011-3

Loftus, E. F., & Pickrell, J. E. (1995). The formation of false memories. *Psychiatric Annals, 25*, 720-725.

Lyle, K. B., & Johnson, M. K. (2006). Importing perceived features into false memories. *Memory, 14*, 197-213. doi: 10.1080/09658210544000060

Maghan, J., O'reilly, G. W., & Shon, P. C. H. (2002). Technology, policing, and implications of in-car videos. *Police Quarterly, 5*, 25-42. doi: 10.1177/109861102129198002

Magnussen, S., & Melinder, A. (2012). What psychologists know and believe about memory: A survey of practitioners. *Applied Cognitive Psychology, 26*, 54-60. doi: 10.1002/acp.1795

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*, 421–428. doi: 10.1016/j.jesp.2007.01.001

Mather, M., Mitchell, K. J., Raye, C. L., Novak, D. L., Greene, E. J., & Johnson, M. K. (2006). Emotional arousal can impair feature binding in working memory. *Journal of Cognitive Neuroscience, 18*, 614-625. doi: 10.1162/jocn.2006.18.4.614

Mazzoni, G. A. L. (2002). Reply to the Commentaries on "Naturally Occurring and Suggestion-Dependent Memory Distortions". *European Psychologist, 7*, 36-38. doi: 10.1027//1016-9040.7.1.36

Mazzoni, G. A. L. (2007). Did you witness demonic possession? A response time analysis of the relationship between event plausibility and autobiographical beliefs. *Psychonomic Bulletin & Review, 14*, 277-281. doi: 10.3758/BF03194064

Mazzoni, G. A. L., & Kirsch, I. (2002). Autobiographical memories and beliefs: A preliminary metacognitive model. In T. Perfect & B. Schwart (Eds.), *Applied Metacognition* (pp. 121-145). Cambridge, England: Cambridge University Press.

Mazzoni, G. A. L., Loftus, E. F., & Kirsch, I. (2001). Changing beliefs about implausible autobiographical events: A little plausibility goes a long way. *Journal of Experimental Psychology: Applied, 7*, 51-59. doi: 10.1037//1076-898X.7.1.51

Mazzoni, G. A. L., & Memon, A. (2003). Imagination can create false autobiographical memories. *Psychological Science, 14*, 186-188. doi: 10.1046/j.1432-1327.1999.00020.x

Mazzoni, G. A. L., Scoboria, A., & Harvey, L. (2010). Nonbelieved memories. *Psychological Science, 21*, 1334 –1340. doi: 10.1177/0956797610379865

McCloskey, M., & Zaragoza, M. (1985). Misleading postevent information and memory for events: Arguments and evidence against memory impairment hypotheses. *Journal of Experimental Psychology: General, 114*, 1-16. doi: 10.1037/0096-3445.114.1.1

McDermott, K. B., & Chan, J. C. K. (2006). Effects of repetition on memory for pragmatic inferences. *Memory & Cognition, 34*, 1273-1284. doi: 10.3758/BF03193271

McNally, R. J., Lasko, N. B., Clancy, S. A., Macklin, M. L., Pitman, R. K., & Orr, S. P. (2004). Psychophysiological responding during script-driven imagery in people reporting abduction by space aliens. *Psychological Science, 15*, 493-497. doi: 10.1111/j.0956-7976.2004.00707.x

Meade, M. L., & Roediger, H. L. (2002). Explorations in the social contagion of memory. *Memory & Cognition, 30,* 995-1009. doi: 10.3758/BF03194318

Meyer, J. R., & Reppucci, N. D. (2007). Police practices and perceptions regarding juvenile interrogation and interrogative suggestibility. *Behavioral Sciences & the Law, 25*, 757-780. doi: 10.1002/bsl.774

Mitchell, K. J., & Zaragoza, M. S. (1996). Repeated exposure to suggestion and false memory: The role of contextual variability. *Journal of Memory and Language, 35*, 246-260. doi: 10.1006/jmla.1996.0014

Morgan, C. A., Southwick, S., Steffian, G., Hazlett, G. A., & Loftus, E. F. (2013). Misinformation can influence memory for recently experienced, highly stressful events. *International Journal of Law and Psychiatry, 36,* 11–17. doi: 10.1016/j.ijlp.2012.11.002

Morris, M. W., & Moore, P. C. (2000). The lessons we (don't) learn: Counterfactual thinking and organizational accountability after a close call. *Administrative Science Quarterly, 45*, 737-765. doi: 10.2307/2667018

Moston, S., & Engelberg, T. (1993). Police questioning techniques in tape recorded interviews with criminal suspects. *Policing and Society: An International Journal of Research and Policy, 3*, 223-237. doi: 10.1080/10439463.1993.9964670

Nash, R. A., & Wade, K. A. (2009). Innocent but proven guilty: Eliciting internalised false confessions using doctored-video evidence. *Applied Cognitive Psychology, 23*, 624-637. doi: 10.1002/acp.1500

Nash, R. A., Wade, K. A., & Brewer, R. J. (2009). Why do doctored images distort memory? *Consciousness and Cognition, 18*, 773-780. doi: 10.1016/j.concog.2009.04.011

Nash, R. A., Wade, K. A., & Lindsay, D. S. (2009). Digitally manipulating memory: Effects of doctored videos and imagination in distorting beliefs and memories. *Memory & Cognition, 37*, 414-424. doi: 10.3758/MC.37.4.414

Newman, E. J., Berkowitz, S. R., Nelson, K. J., Garry, M., & Loftus, E. F. (2011). Attitudes about memory dampening drugs depend on context and country. *Applied Cognitive Psychology, 25*, 675-581. doi: 10.1002/acp.1740

Newman, E. J., Garry, M., Bernstein, D. M., Kantner, J., & Lindsay, D. S. (2012). Nonprobative photographs (or words) inflate truthiness. *Psychonomic Bulletin & Review, 19*, 969–974 doi: 10.3758/s13423-012-0292-0

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*, 175-220. doi: 10.1037/1089-2680.2.2.175

Nisbett, R. E., Zukier, H., & Lemley, R. E. (1981). The dilution effect: Nondiagnostic information weakens the implications of diagnostic information. *Cognitive Psychology, 13*, 248-277. doi: 10.1016/0010-0285(81)90010-4

NPPA. (2013). *NPPA Code of Ethics.* Retrieved from https://nppa.org/code_of_ethics

Odegard, T., & Lampinen, J. (2004). Memory conjunction errors for autobiographical events: More than just familiarity. *Memory, 12*, 288-300. doi: 10.1080/09658210244000621

Oppenheimer, D. M. (2006). Consequences of erudite vernacular utilized irrespective of necessity: problems with using long words needlessly. *Applied Cognitive Psychology, 20*, 139-156. doi: 10.1002/acp.1178

Ost, J., Vrij, A., Costall, A., & Bull, R. (2002). Crashing memories and reality monitoring: Distinguishing between perceptions, imaginations and 'false memories'. *Applied Cognitive Psychology, 16*, 125-134. doi: 10.1002/acp.779

Otgaar, H., Alberts, H., & Cuppens, L. (2012). Ego depletion results in an increase in spontaneous false memories. *Consciousness and Cognition, 21*, 1673–1680. doi: 10.1016/j.concog.2012.09.006

Otgaar, H., Candel, I., Merckelbach, H., & Wade, K. A. (2009). Abducted by a UFO: Prevalence information affects young children's false memories for

an implausible event. *Applied Cognitive Psychology, 23,* 115-125. doi: 10.1002/acp.1445

Otgaar, H., Peters, M., & Howe, M. L. (2012). Dividing attention lowers children's but increases adults' false memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38,* 204-210. doi: 10.1037/a0025160

Perillo, J., & Kassin, S. M. (2011). Inside Interrogation: The lie, the bluff, and false confessions. *Law and Human Behavior, 35,* 327-337. doi: 10.1007/s10979-010-9244-2

Pizarro, D. A., Laney, C., Morris, E. K., & Loftus, E. F. (2006). Ripple effects in memory: Judgments of moral blame can distort memory for events. *Memory & Cognition, 34,* 550-555. doi: 10.3758/BF03193578

Police and Criminal Evidence (PACE) Act (1984)

Poole, D. A., Lindsay, D. S., Memon, A., & Bull, R. (1995). Psychotherapy and the recovery of memories of childhood sexual abuse: U.S. and British practitioners' opinions, practices, and experiences. *Journal of Consulting and Clinical Psychology 63,* 426-437 doi: 10.1037/0022-006X.63.3.426

Rabelais, F. (2007). *Gargantua and Pantagruel*. London, UK: Penguin Books Ltd.

Rachman, S. (2002). A cognitive theory of compulsive checking. *Behavior Research and Therapy, 40,* 624-639. doi: 10.1016/S0005-7967(01)00028-6

Ramsey, R. J., & Frank, J. (2007). Wrongful conviction: Perceptions of criminal justice professionals regarding the frequency of wrongful conviction and

the extent of system errors. *Crime and Delinquency, 53*, 436-470. doi: 10.1177/0011128706286554

Reaves, S. (1987). Digital retouching: Is there a place for it in newspaper photography? *Journal of Mass Media Ethics: Exploring Questions of Media Morality, 2*, 40-48. doi: 10.1080/08900528709358293

Reaves, S. (1991). Digital alteration of photographs in consumer magazines. *Journal of Mass Media Ethics: Exploring Questions of Media Morality, 6*, 175-181. doi: 10.1207/s15327728jmme0603_5

Reaves, S., Hitchon, J. B., Park, S. Y., & Yun, G. W. (2004). If looks could kill: Digital manipulation of fashion models. *Journal of Mass Media Ethics, 19*, 56-71. doi: 10.1207/s15327728jmme1901_5

Redlich, A. D., & Goodman, G. S. (2003). Taking responsibility for an act not committed: The influence of age and suggestibility. *Law and Human Behavior, 27*, 141-156. doi: 10.1023/A:1022543012851

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*, 803-814. doi: 10.1037/0278-7393.21.4.803

Roediger, H. L., Meade, M. L., & Bergman, E. T. (2001). Social contagion of memory. *Psychonomic Bulletin & Review, 8*, 365-371. doi: 10.3758/BF03196174

Rubin, D. C. (2006). The Basic-Systems Model of episodic memory. *Perspectives on Psychological Science, 1*, 277-311. doi: 10.1111/j.1745-6916.2006.00017.x

Rubin, D. C., Schrauf, R. W., & Greenberg, D. L. (2003). Belief and recollection of autobiographical memories. *Memory & Cognition, 31,* 887-901. doi: 10.3758/BF03196443

Rubin, D. C., & Siegler, I. C. (2004). Facets of personality and the phenomenology of autobiographical memory. *Applied Cognitive Psychology, 18,* 913–930. doi: 10.1002/acp.1038

Russano, M. B., Meissner, C. A., Narchet, F. M., & Kassin, S. M. (2005). Investigating true and false confessions within a novel experimental paradigm. *Psychological Science, 16,* 481-486. doi: 10.1111/j.0956-7976.2005.01560.x

Russano, M. B., & Narchet, F. M. (2009). *Who's telling the truth? Police officers' perceptions of lying and truth-telling*. Paper presented at the *Annual meeting of the American Society of Criminology (ASC), Los Angeles Convention Center, Los Angeles, CA.*

Sacchi, D. L. M., Agnoli, F., & Loftus, E. F. (2007). Changing history: Doctored photographs affect memory for past public events. *Applied Cognitive Psychology, 21,* 1005-1022. doi: 10.1002/acp.1394

Sanders, A., & Young, R. (2007). *Criminal justice* (3rd ed.). Oxford, England: Oxford University Press.

Sandham, A., Ormerod, T., Dando, C., Bull, R., Jackson, M., & Goulding, J. (2011). *Scent trails: Countering terrorism through informed surveillance.* Paper presented at the 9th international conference on Engineering psychology and cognitive ergonomics, Orlando, FL.

Schwartz, M. (1982). Repetition and Rated Truth Value of Statements. *The American Journal of Psychology, 95*, 393-407. doi: 10.2307/1422132

Scoboria, A., Mazzoni, G. A. L., & Jarry, J. L. (2008). Suggesting childhood food illness results in reduced eating behavior. *Acta Psychologica, 128*, 304-309. doi: 10.1016/j.actpsy.2008.03.002

Scoboria, A., Mazzoni, G. A. L., Kirsch, I., & Jimenez, S. (2006). The effects of prevalence and script information on plausibility, belief, and memory of autobiographical events. *Applied Cognitive Psychology, 20*, 1049–1064. doi: 10.1002/acp.1240

Scoboria, A., Mazzoni, G. A. L., Kirsch, I., & Relyea, M. (2004). Plausibility and belief in autobiographical memory. *Applied Cognitive Psychology, 18*, 791-807. doi: 10.1002/acp.1062

Scoboria, A., Wysman, L., & Otgaar, H. (2012). Credible suggestions affect false autobiographical beliefs *Memory, 20*, 429-442. doi: 10.1080/09658211.2012.677449

Shanteau, J. (1975). Averaging versus multiplying combination rules of inference judgment. *Acta Psychologica, 39*, 83-89. doi: 10.1016/0001-6918(75)90023-2

Shawyer, A., Milne, R., & Bull, R. (2009). Investigative interviewing in the UK. In T. Williamson, R. Milne & S. Savage (Eds.), *International developments in investigative interviewing*. Cullompton, England: Willan Publishing.

Shu, L. L., Gino, F., & Bazerman, M. H. (2011). Dishonest deed, clear conscience: When cheating leads to moral disengagement and motivated

forgetting. *Personality and Social Psychology Bulletin 37*, 330 –349. doi: 10.1177/0146167211398138

Smeets, T., Merckelbach, H., Horselenberg, R., & Jelicic, M. (2005). Trying to recollect past events: Confidence, beliefs, and memories *Clinical Psychology Review, 25*, 917-934. doi: 10.1016/j.cpr.2005.03.005

Sorochinski, M., Hartwig, M., Osborne, J., Wilkins, E., Marsh, J., Kazakov, D., & Granhag, P. A. (2013). Interviewing to detect deception: When to disclose the evidence? *Journal of Police and Criminal Psychology*. doi: 10.1007/s11896-013-9121-2

Soukara, S., Bull, R., & Vrij, A. (2002). Police detectives' aims regarding their interviews with suspects. *International Journal of Police Science and Management, 4*, 101-114.

Soukara, S., Bull, R., Vrij, A., Turner, M., & Cherryman, J. (2009). What really happens in police interviews of suspects? Tactics and confessions. *Psychology, Crime & Law, 15*, 493-506. doi: 10.1080/10683160802201827

Strange, D., Garry, M., Bernstein, D. M., & Lindsay, D. S. (2011). Photographs cause false memories for the news. *Acta Psychologica, 136*, 90-94. doi: 10.1016/j.actpsy.2010.10.006

Strange, D., Hayne, H., & Garry, M. (2007). A photo, a suggestion, a false memory. *Applied Cognitive Psychology, 22*, 587-603. doi: 10.1002/acp.1390

Thomas, A. K., & Loftus, E. F. (2002). Creating bizarre false memories through imagination *Memory & Cognition, 30*, 423-431 doi: 10.3758/BF03194942

Thomsen, S. R., McCoy, J. K., & Williams, M. (2001). Internalizing the impossible: Anorexic outpatients' experiences with women's beauty and fashion magazines. *Eating Disorders: The Journal of Treatment & Prevention, 9*, 49-64. doi: 10.1080/106402601300187731

Troyer, A. K., Winocur, G., Craik, F. I. M., & Moscovitch, M. (1999). Source memory and divided attention: Reciprocal costs to primary and secondary tasks. *Neuropsychology, 13*, 467-474. doi: 10.1037/0894-4105.13.4.467

Tuna, S., & Freitas, E. (2012). Gendered adverts: an analysis of female and male images in contemporary perfume ads. *Comunicação e Sociedade, 21*, 95-107.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124-1131. doi: 10.1126/science.185.4157.1124

van Bergen, S., Horselenberg, R., Merckelbach, H., Jelicic, M., & Beckers, R. (2010). Memory distrust and acceptance of misinformation. *Applied Cognitive Psychology, 24*, 885–896. doi: 10.1002/acp.1595

van Bergen, S., Jelicic, M., & Merckelbach, H. (2008). Interrogation techniques and memory distrust. *Psychology, Crime & Law, 14*, 425-434. doi: 10.1080/10683160701822533

Wade, K. A., & Garry, M. (2005). Strategies for verifying false autobiographical memories. *American Journal of Psychology, 118*, 587-602.

Wade, K. A., Garry, M., Nash, R. A., & Harper, D. (2010). Anchoring effects in the development of false childhood memories. *Psychonomic Bulletin & Review, 17*, 66-72. doi: 10.3758/PBR.17.1.66

Wade, K. A., Garry, M., Read, J. D., & Lindsay, D. S. (2002). A picture is worth a thousand lies: Using false photographs to create false childhood memories. *Psychonomic Bulletin & Review, 9*, 597-603. doi: 10.3758/BF03196318

Wade, K. A., Green, S. L., & Nash, R. A. (2010). Can fabricated evidence induce false eyewitness testimony? *Applied Cognitive Psychology, 24*, 899–908. doi: 10.1002/acp.1607

Walkley, J. (1987). *Police interrogation*. London, UK: Police Review Publishing.

Wallace, D. B., & Kassin, S. M. (2012). Harmless error analysis: How do judges respond to confession errors? *Law and Human Behavior, 36*, 151-157. doi: 10.1007/s10979-010-9262-0

Warren, A. R., & Lane, P. (1995). Effects of timing and type of questioning on eyewitness accuracy and suggestibility *Memory and testimony in the child witness* (pp. 44-60): Thousand Oaks, CA, US: Sage Publications, Inc.

Weaver, K., Garcia, S. M., Schwarz, N., & Miller, D. T. (2007). Inferring the popularity of an opinion from its familiarity: A repetitive voice can sound like a chorus. *Journal of Personality and Social Psychology, 92*, 821-833. doi: 10.1037/0022-3514.92.5.821

Weinstein, Y., & Nash, R. A. (2013). False recognition of objects in visual scenes: Findings from a combined direct and indirect memory test. *Memory & Cognition, 41*, 60-68. doi: 10.3758/s13421-012-0242-0

Wells, T., & Leo, R. A. (2008). *The wrong guys: Murder, false confessions, and the Norfolk Four*. New York, NY: The New Press.

White, W. S. (2003). Confessions in capital cases. *University of Illinois Law Review*, 979-1036.

Whittlesea, B. W. A. (1993). Illusions of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 1235-1253. doi: 10.1037/0278-7393.19.6.1235

Wilde, O. (1998). *The Picture of Dorian Gray* (N. Page Ed.). Peterborough, CA: Broadview Press Ltd.

Winkielman, P., Schwarz, N., Reber, R., & Fazendeiro, T. A. (2003). Cognitive and affective consequences of visual fluency: When seeing is easy on the mind *Persuasive imagery: A consumer response perspective* (pp. 75-89). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Wixted, J. T., & Stretch, V. (2004). In defense of the signal detection interpretation of remember/know judgments. *Psychonomic Bulletin & Review, 11,* 616-641. doi: 10.3758/BF03196616

Wright, D. B., Memon, A., Skagerberg, E. M., & Gabbert, F. (2009). When eyewitnesses talk. *Current Directions in Psychological Science, 18*, 174-178. doi: 10.1111/j.1467-8721.2009.01631.x

Wright, S. (2005). *I have a pony* [CD]. US: Flashback Records.

Zalman, M., & Smith, B. W. (2007). The attitudes of police executives toward Miranda and interrogation policies. *Journal of Criminal Law and Criminology, 97,* 873-942.

Zaragoza, M. S., & Mitchell, K. J. (1996). Repeated exposure to suggestion and the creation of false memories. *Psychological Science, 7*, 294-300. doi: 10.1111/j.1467-9280.1996.tb00377.x

Zaragoza, M. S., Mitchell, K. J., Payment, K., & Drivdahl, S. (2011). False memories for suggestions: The impact of conceptual elaboration. *Journal of Memory and Language, 64*, 18-31. doi: 10.1016/j.jml.2010.09.004

Zaragoza, M. S., Payment, K. E., Ackil, J. K., Drivdahl, S. B., & Beck, M. (2001). Interviewing witnesses: Forced confabulation and confirmatory feedback increase false memories. *Psychological Science, 12*, 473-477. doi: 10.1111/1467-9280.00388

Zhang, M. (2011, July 27). Julia Roberts makeup ads banned in UK for too much Photoshop. Retrieved from http://www.petapixel.com/2011/07/27/julia-roberts-makeup-ads-banned-in-uk-for-too-much-photoshop/

Zhu, B., Chen, C., Loftus, E. F., He, Q., Chen, C., Lei, X., . . . Dong, Q. (2012). Brief exposure to misinformation can lead to long-term false memories. *Applied Cognitive Psychology, 26*, 301-307. doi: 10.1002/acp.1825

Zhu, B., Chen, C., Loftus, E. F., Lin, C., & Dong, Q. (2013). The relationship between DRM and misinformation false memories. *Memory & Cognition*. doi: 10.3758/s13421-013-0300-2

Appendices

A. Final interview as used in Experiment 4

This feedback form is all about your opinion on the software. It will help us to improve it in later versions, so please be as honest as you can. Please briefly describe what you think this experiment was investigating.

_____

_____

_____

For each of the following statements, you first need to indicate your agreement by answering yes/no. You then need to rate each statement on how much believe it to be true in your opinion, from 1 (Strongly disbelieve) to 9 (Strongly believe). Some questions also require you to describe and/or explain your response when asked. The form is divided into subheadings, referring to different parts of the test.

The software in general

1. Hazard perception tests like this are helpful to learner drivers Yes/No

Strongly disbelieve *Strongly believe*

1 2 3 4 5 6 7 8 9

The practice clip

2. The practice clip was helpful/useful for the rest of the test Yes/No

Strongly disbelieve *Strongly believe*

1 2 3 4 5 6 7 8 9

The bonus clip

3. The incentive motivated me to do well on the bonus clip Yes/No

Strongly disbelieve *Strongly believe*

1 2 3 4 5 6 7 8 9

4. I cheated on the bonus clip Yes/No

Strongly disbelieve *Strongly believe*

1 2 3 4 5 6 7 8 9

If Yes: Remember/Know/Unsure

Please describe/explain what happened

_____

_____

_____

The other clips (excluding the practice and bonus clips)

5. I performed consistently across clips Yes/No

Strongly disbelieve *Strongly believe*

1 2 3 4 5 6 7 8 9

6. I cheated on one or more of the clips Yes/No

Strongly disbelieve *Strongly believe*

1 2 3 4 5 6 7 8 9

Clip number(s)

_____

If Yes: Remember/Know/Unsure

Please describe/explain what happened

_____

_____

_____

7. Would you be interested in taking part in future similar experiments?

    Yes/No

B. Digital editing questionnaire

This questionnaire is about your usage of digital editing software. "Digital editing" can mean anything from electronically cropping a photograph, to professionally cutting and editing videos. Please circle your answers (selecting more than one option if appropriate) and elaborate where necessary.

1. How often do you use digital editing software?

   Never Rarely Sometimes Often All the time

   If 'Never', please go to Question 6

2. What do you use the software for?

   Work Leisure Both

3. What do you edit?

   Videos Photographs Other _____

4. What software do you use? (e.g., Adobe Photoshop, Microsoft Office Picture Manager)

   _____

   _____

   _____

5. What level of editing expertise would you say you have?

   Novice Intermediate Expert

6. When you come across a material in everyday life (e.g., a photograph in a magazine, an advertisement on TV, a video on youtube) do you consider the fact that it might have been digitally altered?

   Never Rarely Sometimes Often All the time

7. In general, do you trust materials that you think have been digitally altered?

   Never Rarely Sometimes Often All the time

8. Has this experiment made you more aware of digital editing?

   Yes / No

9. Do you think you will evaluate digital materials more carefully in the future because of this experiment?

   Yes / No