Believing is seeing

Believing is Seeing:

Cognitive-behavioral Consequences of Belief and Recollection

Jianqin Wang

ISBN: 978-94-9301-482-4

Printed by: Gildeprint – www.gildeprint.nl

MIX
Paper from responsible sources
FSC® C012854
www.fsc.org

BELIEVING IS SEEING:
Cognitive-behavioral Consequences of Belief and Recollection

Dissertation

to obtain the degree of Doctor at the Maastrich University,
on the authority of the Rector Magnificus, Prof.dr. Rianne M. Letschert
in accordance with the decision of the Board of Deans,
to be defended in public on
Wednesday, 10th October 2018, at 10.00 hours

by

Jianqin Wang

Supervisors
Dr. T. Smeets
Prof. dr. H.L.G.J. Merckelbach
Prof. dr. M.L. Howe (City, University of London)

Co-supervisor
Prof. dr. H. Otgaar (Maastricht University & KU Leuven)

Assessment Committee
Prof. dr. M. Jelicic (Chair)
Prof. dr. A. d'Argembeau (Université de Liège)
Prof. dr. P. Santtila (New York University)
Dr. M. Sauerland
Prof. dr. P. de Weerd

TABLE OF CONTENTS

CHAPTER 1

General Introduction

This Chapter is an adapted version of the following paper:

Wang, J., Otgaar, H., Howe, M. L., Smeets, T., & Merckelbach, H. (2015). Can you nonbelieve it: What happens when you do not believe in your memories? *In-Mind*, 11.

Humans act based on their memories. This may sound like a cliché, but the practical meaning of this is illustrated by, for example, the devastating effects of mistaken eyewitness identifications (Wells & Olson, 2003), the far-reaching consequences of innocents who falsely confess to crimes they never committed (Kassin & Gudjonsson, 2004), and the tragedy of adults who erroneously come to believe that they recovered very early memories of abuse experiences (Loftus, 1993). As for these alleged recovered memories: there was a fierce debate among psychologists, therapists, and legal professionals in the 1990s (Howe & Knott, 2015) about their authenticity, a debate that was even characterized as the "memory wars" (Patihis, Ho, Tingen, Lilienfeld, & Loftus, 2014). A recent study has shown that the memory wars have not vanished in the 21st century (Patihis & Pendergrast, in press). It took place against a background of hundreds of lawsuits by adults against their parents because of the alleged abuse memories from childhood that had been "recovered" during therapy in adulthood (Lipton, 1999). Some therapists and clinical psychologists argued that such recovered memories are essentially correct and surface after repression has been lifted due to therapeutic interventions (e.g., hypnosis, sedating drugs). Many researchers, however, contended that recovered memories might, in fact, be false memories produced by risky techniques such as hypnosis and guided imagery (Howe & Knott, 2015; Lambert & Lilienfeld, 2007).

One group of individuals who are particularly interesting in this debate are those who previously claimed to have recovered a memory of a sexually abusive episode, but later retracted those claims (Ost, Costall, & Bull, 2002; Ost, 2017). Do these retractors still have "memories" of these abusive episodes? Or how do these retractors view their earlier experiences? A recent study found that some of these retractors continued to have vivid recollections of an event although they no longer believed that the event actually happened, a phenomenon named *nonbelieved memory* (Ost, 2017). Briefly, this phenomenon has revealed that under certain circumstances, people form memories of events but develop doubts about whether the events have actually occurred. Under these circumstances, people still report having vivid images and recollections of an event and even describe re-living the event, but they do not believe that the event actually happened.

NONBELIEVED MEMORIES

This nonbelieved memory phenomenon was previously assumed to be an extraordinary rare phenomenon, and not until recently has it attracted researchers' attention (Mazzoni, Scoboria, & Harvey, 2010; Otgaar, Scoboria, & Mazzoni, 2014). Below, I will first explain the phenomenon of nonbelieved memories, what factors lead to nonbelieved memories, and describe the methods that have been used to experimentally induce them in the lab. Then, I will show how this phenomenon is relevant to theoretical constructs of memory and their possible cognitive-behavioral consequences.

Recalling Being Kidnapped and Bombed, but not Believing it

Memory scholars agree that in most cases, having a memory implies believing in that specific memory (Scoboria, Mazzoni, Kirsch, & Relyea, 2004). That is, we tend to believe that our recollections of certain events are true and that they actually happened. Yet for some of the events we believe in, we are unable to find associated recollections (e.g., we obviously believe that we were born although we do not have a recollection of being born). Sometimes, our long-held beliefs may be refuted by 'clear evidence', such as a photo displaying how an event really unfolded, which then may result in a no longer believed (i.e., nonbelieved) memory.

For some time, scholars assumed that nonbelieved memories are a rare phenomenon. Anecdotal descriptions of such memories have occasionally surfaced in the literature. For example, the famous developmental psychologist Jean Piaget had a vivid memory of a man attempting to kidnap him when he was two years old. He remembered and described the event in great detail including information that the perpetrator scratched his nurse's face (Piaget, 1951). However, not until thirteen years later, Piaget's former nurse confessed that it was she who had fabricated the event and fed the story to him. Piaget no longer believed that he was almost kidnapped as a child, but he could not stop having vivid visualizations and images of the fabricated kidnapping, as if it had occurred. Oliver Sacks (2005) also described a vivid memory of a bomb that fell close to his home when he was a child, but later he did not believe in his memory because he learned that he was away from home at that time.

Nonbelieved memories are fascinating, if only because the bulk of studies focusing on memory examine and measure *believed memories:* memories of events for which people also strongly believe that the event occurred (Scoboria, Mazzoni, Kirsch, & Relyea, 2004). This is a comprehensible situation as believed memories of events might be seen as the *status quo* of our memory repertoire. While various ideas about memory suggest that memories are typically believed to be true (e.g., James, 1890/1950; Brewer, 1996), it is only recently that memory researchers have started to investigate the possibility that memory might exist without accompanying belief.

Are Nonbelieved Memories Really Rare?

In the first systematic study looking into nonbelieved memories, the frequency and characteristics of nonbelieved memories were surveyed (Mazzoni, Scoboria, & Harvey, 2010). Participants were asked to report a believed memory, a nonbelieved memory, and an event that they believed but did not remember. Almost 25% of the 1593 participants reported having experienced a nonbelieved memory. For example, one participant recollected that he had seen a dinosaur although the belief in the event had vanished. Another person reported a childhood memory of a car accident, but many years later discovered that it actually happened to his brother. Memory characteristics, such as visual details, of nonbelieved memories were also examined. Nonbelieved memories did not differ from believed memories in terms of visual characteristics, clarity, richness, and feeling of reliving, which may explain why nonbelieved memories “feel” so authentic.

Some researchers argued that studies like the one of Mazzoni and colleagues (2010) rely on *directly* asking participants about a typical experience and in this way reveal the purpose of the inquiry. If participants know what the researcher is interested in, then that can artificially inflate the rate of nonbelieved memories. In order to understand the nature and frequency of nonbelieved memories in everyday autobiographical memory, Scoboria and Talarico (2013) used an indirect cueing method, without artificially drawing participants’ attention to nonbelieved memories. Participants were asked to recall events from different ages and then rated the degree of event recollections (memory) and belief in the occurrence of events (belief). With this indirect cueing procedure, only 3% to 6% of the events recalled were nonbelieved memories, a much lower rate than the 25% reported by Mazzoni et al. (2010).

What might explain the discrepancy in frequency of nonbelieved memories elicited by direct and indirect cueing methods? There may be no real discrepancy. Although 25% of people have salient nonbelieved memories when they are interviewed systematically and directly about their nonbelieved memories, the accessibility rate to nonbelieved memories in daily retrospection might be quite low (3-6%), and understandably so: why bother about something that you do not believe. Indeed, a related and interesting question is: how do nonbelieved memories affect people's attitudes and behavior? For instance, will nonbelieved memories of sexual abuse influence retractors' attitudes toward the family members that they previously accused of the abuse? Or will their vivid nonbelieved memories hinder the attempts of retractors to restore their relationship with their families?

Factors that Impact Belief of a Memory

There are many different reasons why people stop believing in their memories, considering the prevalence of nonbelieved memories. A major reason why individuals withdraw their belief for a memory is social feedback such as being told by someone that the remembered event did not happen, as illustrated in the following case (achieved from http://www.falsememoryarchive.com/):

"I remember having chickenpox when I was little. It was horrible, itching, I just hated it. My parents treated it with some white solution that came in a brown glass bottle; I remember exactly how it looked. Also I remember going to the doctor and looking at all the red itching spots I had. My mother told me that it never happened, I never actually had chickenpox."

Table 1.1
Reasons that people retract beliefs in events (adapted from Scoboria, Boucher, & Mazzoni, 2014)

Reasons to retract belief	Percent of participants; Primary reason	Percent of participants who mentioned it
Social feedback	42.2%	52.8%
Event plausibility	19.5%	35.4%
Alternative attributions (dreams, etc.)	8.8%	28.9%
Internal features of event representations	7.2%	16.3%
External evidence (e.g., photos)	7.2%	10.7%
Notions of self/others	6.4%	12.0%
Personal motivation	1.1%	4.3%

A recent study has shown that 42% percent (*n* = 158) of the people with nonbelieved memories indicated social feedback as their primary reason to relinquish their belief (Scoboria, Boucher, & Mazzoni, 2014). Indeed, we are not unfamiliar with this scene: two friends fight with each other about whose memory is the correct version of what has happened, until one is persuaded by the other. Or twins dispute the ownership of a memory that both recollected being pushed off a bike by a cousin, although only one was pushed off a bike (Sheen, Kemp, & Rubin, 2001, 2006). However, when it comes to two witnesses in criminal investigations, the consequence is much more serious. For instance, one witness might be informed by the police officer about what another witness had said (Luus & Wells, 1994). When the information provided by the police officer contradicts with the witness's memory, the witness may undermine his[1] belief in his memory and choose to not report certain details.

The second major factor that could lead to nonbelieved memories is event plausibility, referenced by 19.5% of the sample (*n* = 73). For example, one individual reported that "*I remember running away from the hospital as a new born baby*", which the person knew was impossible though the memory felt clear and vivid. Other examples include remembering seeing Santa Claus or memories of flying.

Internal features of event representations are also a frequently referenced reason for withdrawing belief. Features of the internal memory representations might lead individuals to question the memory. Some retractors of sexual abuse memories reported that they had come to doubt the 'truth' of their claims because of the experiential qualities of the recovered memories themselves (Ost et al., 2002). They stated that the memories were sometimes not familiar or lacked context, or were even too clear to be true. As one respondent said: "*The recovered memories were crystal clear in every detail (that's how I realized they were not true) ... my real memories faded with time unless I saw photos or talk with relatives*".

Another factor worth noting is personal motivation where individuals express a desire to not remember the event. In Scoboria et al.'s (2014) research, there is a certain percentage of people who were uncomfortable with or disliked the content of their memories and who "pushed it away from my mind" or "did not want to believe that that happened". They

[1] Here and elsewhere, "he" and "his" may also refer to "she" and "her".

successfully compelled themselves to withdraw belief from their memories, thus forming nonbelieved memories. Although only 4.3% of the respondents mentioned personal motivation as a factor leading to nonbelieved memories, considering that the sample were college students, the frequency of nonbelieved memories in populations with negative or traumatic experiences is probably much higher.

Creating Nonbelieved Memories by Providing Feedback

Besides studying naturally occurring nonbelieved memories, researchers have examined whether nonbelieved memories can be elicited in the laboratory. Although a plethora of research has revealed that repeatedly suggesting that a false event occurred increases confidence, belief, and even recollection for that event (e.g., Garry & Wade, 2005; Koehler, 1991; Mazzoni, Loftus, & Kirsch, 2001; Mazzoni & Memon, 2003), our knowledge about how to undermine belief is still quite limited. In one demonstration, Clark, Nash, Fincham, and Mazzoni (2012) provided participants with fake videos (i.e., doctored-video paradigm; Nash, Wade, & Lindsay, 2009) to create nonbelieved memories. Participants were first asked to perform some actions such as clapping their hands and rubbing the table while their actions were being video recorded. Two days later, participants watched a doctored video edited by the researchers in which fake actions were embedded and in this way, participants falsely “remembered” and “believed” that they had performed the fake actions. Finally, researchers told them that actually the video clip was doctored and then measured their belief and recollection for the action. This manipulation undermined participants’ belief for 14% to 26% of the fake actions but vivid recollections for these fake actions remained.

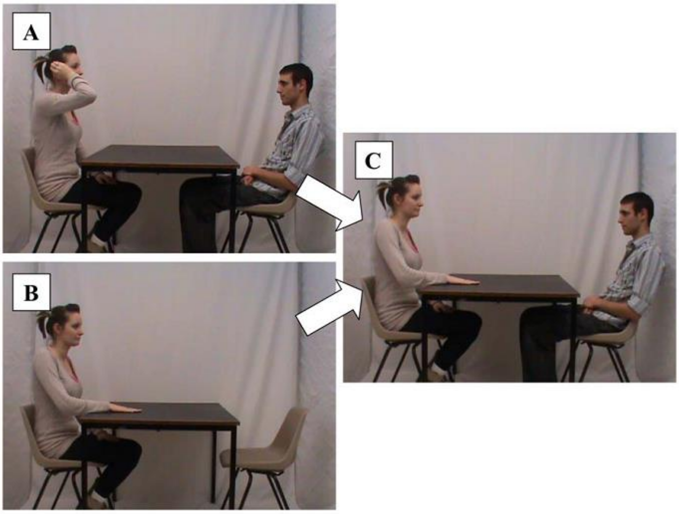


Figure 1.1. Doctored video clip. (A) Real clip. (B) Fake action. (C) Doctored composite of (A) and (B). From Clark, A., Nash, R. A., Fincham, G., & Mazzoni, G. (2012). Creating non-believed memories for recent autobiographical events. *PLoS ONE, 7*(3): e32998. Copyright 2012 by Andrew Clark. Reprinted with permission.

In another demonstration, Otgaar, Scoboria, and Smeets (2013) used a false memory implantation procedure to experimentally elicit nonbelieved memories in children and adults. The false memory implantation method is known as an effective manipulation to create vivid mental representations for suggested false events in both children and adults (see Otgaar, Verschuere, Meijer, & Van Oorsouw, 2012). In general, in this paradigm, individuals are: (1) provided with a fictitious event such as a hot balloon ride when they were children; (2) guided by the experimenter to mentally "travel" back to the suggested event and think about or imagine the details of that experience; and (3) asked whether they have formed a memory for the false event across multiple interviews. A mega-analysis of this memory implantation approach found that averagely 30.6% of the participants would develop false memories for the suggested false events (Scoboria, Wade, Lindsay, Azad, Strange, Ost, & Hyman, 2017).

Otgaar and colleagues (2013) adapted this implantation method to elicit nonbelieved memories by debriefing individuals after the final interview and suggesting that the event did not happen and then asking them whether they still believed or recollected the event. They found that for 45% (n = 29) of children participants who formed false memories, belief for the false event (i.e., hot balloon ride) decreased, but that their memory representations for the hot balloon ride remained vivid.

Not only can people be led to form nonbelieved memories for false events, but they can also be induced to stop believing in memories for true experiences. Mazzoni, Clark, and Nash (2014) used the same procedure—the doctored video paradigm—to examine whether it was possible to persuade people to disbelieve recollected events that actually did happen. This topic is worth investigating because there are important real life equivalents. Consider the child victim who is told by the perpetrator that the abuse did not happen, or witnesses who are told during interrogations that genuine experiences did not occur. Mazzoni and colleagues found that it was possible to undermine autobiographical belief for genuinely performed actions. More recently, Otgaar, Wang, Fränken, and Howe (2018) showed that social feedback coupled with evidence affected participants' beliefs in experiencing virtual reality scenes but did not impact their recollections. Research also found that people did not believe that they performed certain actions after their memories were challenged, although they actually performed the actions (Otgaar, Scoboria, Howe, Moldoveanu, & Smeets, 2016).

THEORETICAL CONSTRUCTS OF MEMORY: BELIEF VERSUS RECOLLECTION

What is memory? More than a century ago, James (1890) proposed that "the mental side [of recall] is the conscious vision of the past occurrence, and the belief that we experienced it before" (p. 655, Chapter 16, *Principles of Psychology*). In his speculation, memory is not only the mental reminiscence, re-perception or *recollection* of a past experience, but is also accompanied by the *belief* that the experience actually occurred. Indeed, for most of the events that we can remember, we also believe that these events happened in the past. Recollection and belief judgments in remembering seem to be entangled with each other. It is only recently that research on nonbelieved memories has suggested that belief and recollection may be two distinct components of autobiographical memory.

The Nested Model of Belief and Memory

Scoboria, Mazzoni, Kirsch, and Relyea (2004) proposed a nested model according to which memory is a construct nested within the belief construct (see Figure 1.2). That is, for memories of events, people also have beliefs that the events occurred. However, for events that they believe happened, people do not necessarily have memories. For example, people believe that they have gone for a balloon ride with their parents when they were young after seeing a photograph, although they could not remember the balloon ride. Thus, the relationship between belief and memory constructs appears to be such that memory is nested within belief. Scoboria and colleagues asked participants to rate memory and belief ratings for ten distinct events such as getting lost, choking on a small object, and getting abducted by an UFO. They found that for all these events with different levels of perceived plausibility, the average belief rating was always higher than the average memory rating, which confirms the nested relationship between belief and memory.

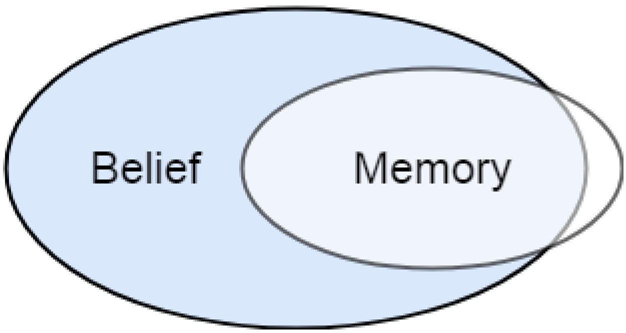


Figure 1.2. The nested model of belief and memory. Adapted from Scoboria, Mazzoni, Kirsch, and Relyea (2004).

Nonetheless, in the nested model, there are a small portion of events for which people have memories (i.e., recollections) but no beliefs (see Figure 1.2, the white portion on the right). Those nonbelieved memories, as we have reviewed before, are occasions where people's memory ratings exceed belief ratings for events. After analysing memory and belief rating data of 685 participants, Scoboria et al. (2004) found that in 96% of a total of 6850 instances, memory was nested within belief, while in the other 4% of the instances, memory was not nested within belief (see Figure 1.2). Thus the nested model implies partial dissociation between belief and recollection.

To be noted, the primary aim of the nested model was to explain the formation of false memories. It is proposed that when people are suggested an event, belief should form before a memory develops (see also Mazzoni, Loftus, & Kirsch, 2001; Pezdek, Finger, & Hodge, 1997). That is, in order to develop a false memory, one needs first to believe in the occurrence in the event.

The Basic-Systems Model of Episodic Memory

Rubin (2005, 2006) proposed a basic-systems model of episodic memory where basic systems such as different sensory systems and coordination systems are integrated as interactive components of episodic memory. Compared to traditional models of memory such as the Modal Model (Atkinson & Shiffrin, 1968), which are often computer metaphors involving abstract propositions and are unimodal, the basic-systems model postulates that episodic memories usually are multimodal such as vision, hearing, smell and body sense, and they require spatial, temporal and even emotional contextual characteristics. Of central importance in the basic-systems model are two key properties of memory: a sense of recollection and the belief that memories are accurate (Rubin, 2006). These two properties are metacognitive judgments based on processing in the basic systems and both constitute the feeling of remembering in the end.

Recollection, according to Rubin (2005), is a sense of re-living a past experience, which distinguishes memory from other mental activities such as imagination. Belief, on the other hand, refers to the extent to which people believe the event really occurred in the way they remembered it (Rubin, 2006). For example, when testifying in court, witnesses are often asked “Where were you when the crime happened?”. This question would cue witnesses to mentally travel back to the time when the crime happened and relive the visual, auditory, spatial, contextual and emotional aspects of the crime. Meanwhile, witnesses also usually believe that what they recall accurately represents what happened when they testify (otherwise they would not testify).

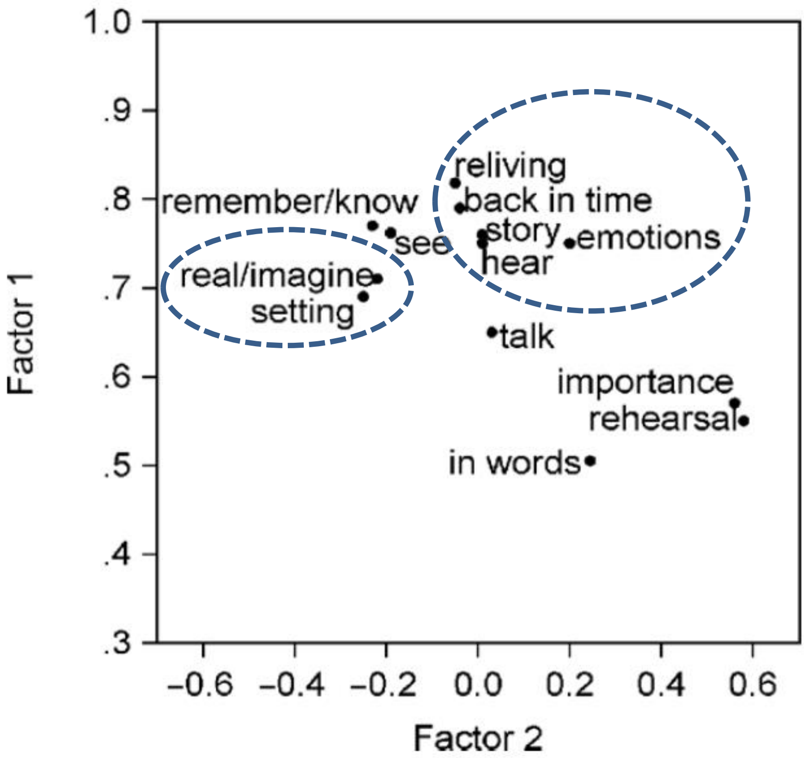


Figure 1.3. Graphical demonstration of variable clusters. “Real/imagine” measures belief in accuracy and “reliving” measures recollection. Note that the closer two variables are, the higher correlation they have. Adapted and retrieved from Rubin, Schrauf, and Greenberg (2003).

In the basic-systems model, belief and recollection are not separate subsystems but rather metacognitive judgments of the basic-systems (Rubin et al., 2003; 2006). The basic systems of memory include vision, spatial imagery, audition, olfaction, emotion, motor outputs and other senses and coordination. In empirical research, the degree of recollection (i.e., the degree to which people felt reliving the original event) was found to be predicted by visual imagery, auditory features, coherent story and emotions (R^2 ~70%), whereas belief in memory accuracy (i.e., the degree to which their memory is real or imagined) was predicted by contextual aspects such as how well the setting was remembered (see dashed circles in Figure 1.3; Rubin, Schrauf, & Greenberg, 2003; Rubin & Siegler, 2004). Hence, the basic-systems model shows the dissociation of recollection and belief in that each is supported by different systems.

The Dissociated Model of Belief and Recollection

In the past several years, as we have reviewed, evidence has accumulated that nonbelieved memories are not rare and are relatively easy to create (e.g., Mazzoni et al., 2011; Clark et al., 2012; Otgaar et al., 2014), suggesting that belief in occurrence of an event is a dissociable component from memory. More recently, Scoboria, Jackson, Talarico, Hanczakowski, Wysman, and Mazzoni (2014) proposed that autobiographical belief and recollection are two independent and continuous components of memory. As Figure 1.4 shows, event representations can be categorized based on two dimensions of memory: high or low in autobiographical belief as well as high or low in recollection. Events for which people have both high recollection and high belief, are represented as believed memories. Scoboria and Talarico (2013) asked participants to recall childhood events and found that the majority of events (58.6%; *n* = 501) that were reported were believed memories. Another 38.3% (*n* = 328) of the recalled events were believed but not remembered events (i.e., events people believe happened but cannot recollect). For events that people elaborate with low belief, there are nonbelieved memories (high recollection) as well as nonbelieved and not remembered events (low recollection).

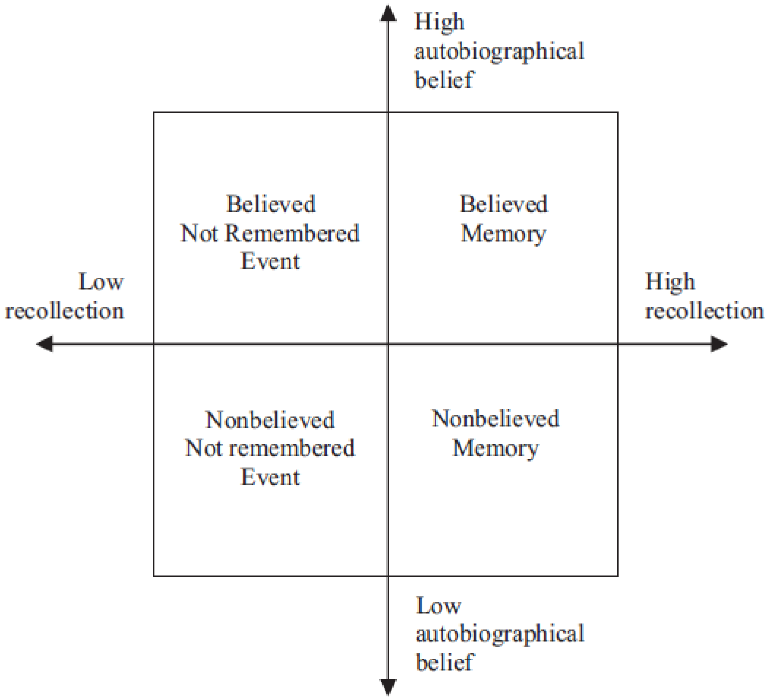


Figure 1.4. The dissociated model of autobiographical belief and recollection. From Scoboria, Jackson, Talarico, Hanczakowski, Wysman, and Mazzoni (2014).

Scoboria et al. (2014) cued around 300 participants (Study 1) to recall various childhood events and asked them to rate the events in terms of belief, recollection, and memory aspects similar to the basic systems of Rubin (2006). Using structural equation modeling, it was clearly shown that recollection and belief were distinct constructs: factors that predicted recollection (including perceptual, re-experiencing, emotion intensity, event specificity) did not predict belief, while belief was predicted strongly by plausibility, which in turn only predicted recollection weakly. The data support the dissociated model of belief and recollection of memory.

To be noted, the belief construct in the dissociated model refers to the truth value attached to the event (i.e., whether it occurred or not) which is conceptually different from "belief in memory accuracy" (i.e., whether memory accurately represents the event) of Rubin's (2006) basic-systems model. However, the concept of "belief in occurrence" is in line with literal definitions of belief such as that of James (1890). Hence, in this dissertation, the term "belief" always refers to people's belief in the occurrence of an event. For the concept of recollection, there is little variation among different models, which refers to mentally re-experiencing or re-perceiving an event.

To summarize, empirical research and theories of memory so far have agreed that belief and recollection are two distinct constructs of memory. However, the issue that arises then is that most memories currently being studied are *believed memories* and research typically does not distinguish belief from recollection (Scoboria et al., 2014). For instance, the Life Event Inventory (LEI; Mazzoni & Loftus, 1996) was intended to measure people's childhood memories but it asked participants to identify which events ever happened to them, which could be either a belief or recollection (Scoboria et al., 2004). Another example is that when trying to study false memories induced by suggestion, researchers simply used the term false memory to refer to both false belief and false memory (see Bernstein & Loftus, 2009; Laney, Morris, Bernstein, Wakefield, & Loftus, 2008). The confounding of belief and recollection is especially relevant when it comes to studying the consequences of memories. So, when memories are found to impact people's cognitive-behavioral performance, which component (belief vs. recollection) truly drives their behavior? This dissertation will revolve around this question concerning the cognitive-behavioral consequences of belief and recollection.

THE CONSEQUENCES OF BELIEF AND RECOLLECTION

There is little doubt that memories (both true and false) can impact behavior. In court, memories of eyewitnesses serve as important and oftentimes the sole evidence of the crime; they can help exert justice when they are genuine representations of the crime or they may lead to miscarriages of justice when memories contain errors (Howe, Knott, & Conway, 2018; Wells & Olson, 2003). In the lab, an abundance of research has demonstrated the priming effect of true and false memories on various tasks (e.g., Howe, Garner, Dewhurst, & Ball, 2010; Tulving & Schacter, 1990). In **Part I** of this dissertation, I focus on the consequences of general memories in practice and in the lab. I review the negative consequences of false memories in the legal arena in Chapter 2 and I examine the positive consequences of true and false memories in Chapter 3. The purpose of the research presented in my dissertation is to show in diverse ways how belief and recollection impact memory-related behavior. Therefore, in **Part II,** I examine the impact of belief and recollection on different types of memory-related behavior. The topics involved in **Part II** are as follows.

Problem Solving

Memories have been found to prime problem solving behavior (Howe, Garner, Dewhurst, & Ball, 2010; Howe, Garner, Charlesworth, & Knott, 2011; Howe, Wilkinson, Garner, & Ball, 2016). In an insight based problem solving task, people are asked to solve a three-word puzzle (e.g., *Board/Mail/Magic*) by coming up with a word that links to all the three words (i.e., *black*) (e.g., Bowden, Jung-Beeman, Fleck, & Kounios, 2005; Howe et al., 2010). If the solution to the puzzle has been presented before or falsely remembered in memory, people normally solve the problems faster than those whose solutions are not primed by memory. With the dissociation of belief and recollection for memory, it is unknown whether the priming effect is due to belief or recollection of memories. In Chapter 4, I examine this question.

Decision Making

In classical conditioning and reward learning research, memories are thought to guide organisms' decisions to avoid harm or gain reward. For example, in a sensory preconditioning task, a dog learns a "whistle – light" association first and learns "light – electric shock" later. The dog would then expect an electric shock after a whistle and flexes its limb to avoid harm, even though the whistle was never paired with the shock (Brogden, 1939). This sensory preconditioning takes place in humans as well. If people first learn a "picture – circle" association and then learn that the circle wins money in a different occasion, they would show a decision bias to choose that picture to win money, even though the picture has never been rewarded (Wimmer & Shohamy, 2012). People make these kinds of biased decisions because they have formed memory associations (e.g., picture – circle – money) that enable the value of money to transfer via the circle to the picture. Hence, if the belief for the memory association (e.g., picture – circle) is undermined, will people's decision making be altered? Chapter 5 tried to undermine participants' beliefs for established memory associations to examine the roles of belief and recollection on decision making.

Food Preferences

Recent work on the behavioral consequences of false memories has shown that false memories of food experiences impact food behavior (for a review see Bernstein & Loftus, 2009). Typically, participants in this type of study complete a food history questionnaire and then are told that based on their answers, the computer has generated a health profile for them. The profile falsely suggests that as a child, they had gotten ill after eating a particular food, such as egg-salad. After about two weeks, their eating behavior is measured. People receiving the suggestion have been shown to consume significantly less egg-salad compared to the control group, and this effect can last for months (Bernstein, Laney, Morris, & Loftus, 2005; Scoboria, Mazzoni, Jarry, & Bernstein, 2012). Chapter 6 dissociated participants' false beliefs and false recollections of a food aversive event, and then investigated whether it was belief or recollection that truly impacted food preferences.

To conclude, a novel line of research has emerged showing that recollection and belief are distinct constructs ending up in the creation of different mental representations of events such as nonbelieved memories. The purpose of **Part II** was to examine the cognitive-behavioral consequences of these two memory constructs. In **Part III**, I studied and discussed mechanisms underlying memory and how they might impact behavior. I have proposed possible mechanisms on how memory impacts behavior and this will be described as well in the General Discussion.

References

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In *Psychology of learning and motivation* (Vol. 2, pp. 89-195). Academic Press.

Bernstein, D. M., Laney, C., Morris, E. K., & Loftus, E. F. (2005). False beliefs about fattening foods can have healthy consequences. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 13724-13731.

Bernstein, D. M., & Loftus, E. F. (2009). The consequences of false memories for food preferences and choices. *Perspectives on Psychological Science*, *4,* 135–139.

Bernstein, D. M., Scoboria, A., & Arnold, R. (2015). The consequences of suggesting false childhood food events. *Acta Psychologica, 156,* 1–7.

Bowden, E.M., Jung-Beeman, M., Fleck, J., & Kounios, J. (2005). New approaches to demystifying insight. *Trends in Cognitive Sciences, 9,* 322–328.

Brewer, W. F. (1996). What is recollective memory? In D. C. Rubin (Ed.), *Remembering our past: Studies in autobiographical memory* (pp. 19-66). Cambridge, England: Cambridge University Press.

Brogden, W. J. (1939). Sensory pre-conditioning. *Journal of Experimental Psychology*, *25*, 323-332.

Clark, A., Nash, R. A., Fincham, G., & Mazzoni, G. (2012). Creating non-believed memories for recent autobiographical events. *PLoS ONE, 7*: e32998. Retrieved from http://www .plosone.org/article/info%3Adoi%2F10.1371%2Fjournal .pone.0032998.

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

Howe, M. L., Garner, S. R., Charlesworth, M., & Knott, L. (2011). A brighter side to memory illusions: False memories prime children's and adults' insight-based problem solving. *Journal of Experimental Child Psychology, 108*, 383–393.

Howe, M. L., Garner, S. R., Dewhurst, S. A., & Ball, L. J. (2010). Can false memories prime problem solutions? *Cognition, 117,* 176–181.

Howe, M. L., & Knott, L. M. (2015). The fallibility of memory in judicial processes: Lessons from the past and their modern consequences. *Memory*, *23*, 633-656.

Howe, M. L., Knott, L. M., & Conway, M. A. (2018). *Memory and Miscarriages of Justice.* Abingdon, UK: Routledge.

Howe, M. L., Wilkinson, S., Garner, S. R., & Ball, L. J. (2016). On the adaptive function of children's and adults' false memories. *Memory*, *24*, 1062–1077.

James, W. (1890/1950). *The principles of psychology* (Vol. I). New York: Dover.

Kassin, S. M., & Gudjonsson, G. H. (2004). The psychology of confession evidence: A review of the literature and issues. *Psychological Science in the Public Interest, 5,* 35–69.

Koehler, D. J. (1991). Explanation, imagination, and confidence in judgment. *Psychological Bulletin*, *110,* 499-519.

Lambert, K., & Lilienfeld, S. (2007). Brain stains. *Scientific American Mind, 18*, 46–53.

Lipton, A. (1999*).* Recovered memories in the courts. In S. Taub (Ed.), *Recovered memories of child sexual abuse: Psychological, social, and legal perspectives on a contemporary mental health controversy* (pp. 165–210). Springfield, IL: Charles C. Thomas.

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.

Loftus, E. F. (1993). The reality of repressed memories. *American Psychologist*, *48*, 518-537

Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning and Memory*, *12*, 361–366.

Luus, C. A. E., & Wells, G. L. (1994). The malleability of eyewitness confidence: Co-witness and perseverance effects. *Journal of Applied Psychology*, 79, 714–723.

Mazzoni, G. A. L., Clark, A., & Nash, R. A. (2014). Disowned recollections: Denying true experiences undermines belief in occurrence but not judgments of remembering. *Acta Psychologica*, *145,* 139–146.

Mazzoni, G. A., & Loftus, E. F. (1996). When dreams become reality. *Consciousness and Cognition*, *5*(4), 442-462.

Mazzoni, G., & Memon, A. (2003). Imagination can create false autobiographical memories. *Psychological Science*, *14*, 186-188.

Mazzoni, G. A., 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.

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

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.

Ost, J., Costall, A., & Bull, R. (2002). A perfect symmetry? A study of retractors' experiences of making and then repudiating claims of early sexual abuse. *Psychology, Crime & Law, 8,* 155–181.

Ost, J. (2017). Adults' retractions of childhood sexual abuse allegations: high-stakes and the (in) validation of recollection. *Memory, 25,* 900-909.

Otgaar, H., Scoboria, A., Howe, M. L., Moldoveanu, G., & Smeets, T. (2016). Challenging memories in children and adults using an imagination inflation procedure. *Psychology of Consciousness: Theory, Research, and Practice, 3,* 270-283.

Otgaar, H., Scoboria, A., & Mazzoni, G. (2014). On the existence and implications of nonbelieved memories. *Current Directions in Psychological Science, 23*, 349–354.

Otgaar, H., Scoboria, A., & Smeets, T. (2013). Experimentally evoking nonbelieved memories for childhood events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 717–730.

Otgaar, H., Verschuere, B., Meijer, E. H., & van Oorsouw, K. (2012). The origin of children's implanted false memories: Memory traces or compliance. *Acta Psychologica, 139,* 397–403.

Patihis, L., Ho, L. Y., Tingen, I. W., Lilienfeld, S. O., & Loftus, E. F. (2014). Are the "memory wars" over? A scientist-practitioner gap in beliefs about repressed memory. *Psychological Science, 25,* 519–530.

Patihis, L., & Pendergrast, M. H. (in press). Reports of recovered memories of abuse in therapy in a large age-representative U.S. national sample: Therapy type and decade comparisons. *Clinical Psychological Science.*

Piaget, J. (1951). *Play, dreams, and imitation in childhood* (L. Gattegno & F.M. Hodsen, Trans.). New York: Norton.

Rubin, D. C. (2005). A basic-systems approach to autobiographical memory. *Current Directions in Psychological Science*, *14*, 79-83.

Rubin, D. C. (2006). The basic-systems model of episodic memory. *Perspectives on Psychological Science*, *1*, 277–311.

Rubin, D. C., Schrauf, R. W., & Greenberg, D. L. (2003). Belief and recollection of autobiographical memories. *Memory & cognition*, *31*, 887-901.

Sacks, O. (2005). On memory. *The Threepenny Review*. Retrieved from http://www.threepennyreview.com/samples/sacks_w05.html

Scoboria, A., Boucher, C., & Mazzoni, G. (2015). Reasons for withdrawing belief in vivid autobiographical memories. *Memory, 23,* 545–562.

Scoboria, A., Jackson, D., Talarico, J., Hanczakowski, M., Wysman, L., & Mazzoni, G. (2014). The role of belief in occurrence within autobiographical memory. *Journal of Experimental Psychology: General, 143,* 1242–1258.

Scoboria, A., Mazzoni, G., Jarry, J. L., & Bernstein, D. M. (2012). Personalized and not general suggestion produces false autobiographical memories and suggestion-consistent behavior. *Acta Psychologica, 139,* 225–232.

Scoboria, A., Mazzoni, G., Kirsch, I., & Relyea, M. (2004). Plausibility and belief in autobiographical memory. *Applied Cognitive Psychology*, *18*, 791–807.

Scoboria, A., & Talarico, J. M. (2013). Indirect cueing elicits distinct types of autobiographical event representations. *Consciousness and Cognition, 22,* 1495–1509.

Scoboria, A., Wade, K. A., Lindsay, D. S., Azad, T., Strange, D., Ost, J., & Hyman, I. E. (2017). A mega-analysis of memory reports from eight peer-reviewed false memory implantation studies. *Memory*, *25*, 146-163.

Sheen, M., Kemp, S., & Rubin, D. (2001). Twins dispute memory ownership: A new false memory phenomenon. *Memory & Cognition*, *29*, 779-788.

Sheen, M., Kemp, S., & Rubin, D. C. (2006). Disputes over memory ownership: What memories are disputed?. *Genes, Brain and Behavior, 5,* 9-13.

Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science*, *247*(4940), 301-306.

Wells, G. L., & Olson, E. A. (2003). Eyewitness testimony. *Annual Review of Psychology*, *54*, 277-295.

Wimmer, G. E., & Shohamy, D. (2012). Preference by association: how memory mechanisms in the hippocampus bias decisions. *Science, 338*, 270-273.

Part I

Consequences of Memories in Practice and in the Lab

CHAPTER 2

Negative Consequences of False Memories

This Chapter is an adapted version of the following paper:

Wang, J., Otgaar, H., Zhou, C., Howe, M. L., Smeets, T., & Merckelbach, H. (2018). Consequences of false memories in eyewitness testimony: A review and implications for the Chinese legal system. *Psychological Research on Urban Society*, *1*, 12-25.

Abstract

False memories can have severe legal consequences, including the imprisonment of innocent people. For example, false eyewitness memory is the largest factor contributing to miscarriages of justice in the United States. To date, no study has focused on how false memories might play a role in the Chinese legal system. The purpose of this review is to summarize the latest findings on false memory and eyewitness testimony in the English literature, and to shed some light on how the Chinese legal system may incorporate these experiences into practice. Overall, false memories of eyewitnesses are generated either by external misleading information or by internal cognitive processes; false memories may guide police investigation in a wrong direction, or cause eyewitnesses to misidentify an innocent person as the perpetrator. We conclude that specially designed interview protocols such as the Cognitive Interview, warnings given to eyewitnesses, and blind lineup administration might prevent or lower the risk of false memory occurrence.

Keywords: false memory, eyewitness testimony, Chinese legal practice

Somewhere in December, 2003, Haisheng Zhang (张海生) was visiting his relatives in Lichuang County, Henan Province, China. Suddenly, he was detained by the police as a suspect for raping a 13-year-old girl in the woods. He was eventually sentenced to 9 years of imprisonment by the Court of Lichuang County. The most important piece of evidence was the testimony by the victim stating that she was completely confident that Zhang was the culprit. Besides the testimony, three other teenage girls identified Zhang from a lineup as the person who talked to the victim and led the victim to the woods. Meanwhile, there was no physical evidence incriminating Zhang as the offender. More than a year after his conviction, another defendant, who was recently caught, confessed to a series of sex offending cases, including the one that Zhang was being charged with. When Zhang was released because of this confession, he had spent 480 days in prison.[2]

This is not the only Chinese case in which an innocent person was falsely convicted and imprisoned because of erroneous memories. Another case occurred in 1990 and was recently revised as well. In this case, Jibin Xu (徐继彬) was wrongfully convicted of rape because he was identified by the victim as the assailant, even though the police should have found out that his blood type did not match the blood type of the real perpetrator. Only after 16 years was he proven innocent by a blood test.[3]

In these cases, innocent people were convicted because of the absolute reliance on eyewitness testimonies even when they included erroneous memories implying someone's guilt. In the absence of physical evidence, these testimonies became crucial. Importantly, in the majority of criminal proceedings, eyewitness testimonies are the most important piece of evidence (e.g., Howe, Knott, & Conway, 2018). Objective evidence such as DNA is frequently lacking (Howe & Knott, 2015; Peterson, Hickman, Strom, & Johnson, 2013). The result of this is that legal professionals often must rely on the memory of a victim and/or witness. However, memory is a flexible system that is not as reliable as people expect (Loftus, 2004; Schacter, 2012). Our memories are fallible. That is, they are not literally reproduced but reconstructed when they are retrieved (e.g., Howe et al., 2018). During such reconstruction, unintentional errors might slip in, which can lead to the occurrence of false memories. False memories refer to memories of events that did not

[2] http://news.sina.com.cn/c/2005-04-14/04075642673s.shtml
[3] http://news.sina.com.cn/c/2007-03-20/120311454076s.shtml

happen, yet are experienced as real (Loftus, 2005). Although such false memories can occur in many different situations (e.g., misremembering that you placed your car keys on the table while in fact they were still in the car) and are oftentimes relatively innocuous, when they enter the legal arena, consequences can be quite dramatic especially when they involve false accusations of sexual abuse or faulty eyewitness identifications (Otgaar, Sauerland, & Petrila, 2013).

In the legal arena, eyewitness misidentification has been shown to be the largest contributing factor leading to wrongful convictions, playing a role in more than 70% (n = 243) of convictions which were later overturned through DNA testing in the United States (data achieved from the Innocent Project, http://www.innocenceproject.org/causes/eyewitness-misidentification/). Data about factors responsible for miscarriages of justice are non-existent in China. The main purpose of this review is to assemble the most recent findings on false memories and eyewitness testimony. First, we will review whether the issue of eyewitness testimony has attracted the attention of scholars and legal professionals in China. Next, we will discuss classical research paradigms that demonstrate the malleability of memory. Following this, we will concentrate on the latest findings in the field of eyewitness false memory, which have mainly been published in English journals. Finally, conclusions and implications for legal practice in China are discussed.

Eyewitness Testimony in Chinese Cases

Chen (2007) reviewed 20 nationally-known wrongly convicted cases that were exposed by the media and concluded that torture, improper evidence collection, and ignoring the use of scientific techniques were the most frequently mentioned risk factors in these cases. One limitation of this review is that the cases described by Chen were “famous” cases exposed by journalists, who were particularly interested in uncovering cases involving torture. There was no focus on whether any erroneous memories might have been present in such cases.

To our knowledge, very few studies have ever focused on the important role that erroneous eyewitness testimonies and hence, false memories, might have played in legal proceedings in China. To examine this issue more closely, we selected the China National

Knowledge Infrastructure (CNKI) Database, which covers 10,267 Chinese academic journals (almost all academic journals in China), and the CNKI Masters' Theses Database and Doctoral Dissertations Database to search for literature on false memories and eyewitness testimony at the time of this writing. The following keywords were selected to search for the literature: Eyewitness, eyewitness testimony, eyewitness memory, false memory, eyewitness events, children witnesses, or memory distortions. This literature search resulted in 18 papers and six theses/dissertations concerning eyewitness memory, seven papers on eyewitness identification, and nine papers on child witnesses, dated from 1991 to 2016. After reviewing these papers, we found that not one paper specifically looked at the relationship between legal cases and false memories. This suggests that in the Chinese psychological literature, the topic of false memories in the courtroom has not attracted much attention, although we know from many cases in other countries that false memories constitute an important source of wrongful convictions (Garrett, 2011; Loftus, 2013). Based on this observation and the Chinese cases reviewed earlier, it is likely that false memories might have affected testimonies in Chinese cases as well.

The Fallibility of Memory

The idea that memory can be easily distorted has been examined by relying on false memory paradigms in which false memories are produced suggestively or spontaneously. Loftus (1975) first demonstrated how leading questions could impact eyewitness reports by using the *misinformation paradigm*. Basically, the misinformation paradigm consists of three stages. First, in the encoding phase, participants generally view a video depicting an event such as a crime or an accident. Then participants are exposed to misleading information (e.g., in the form of statements or leading questions). Finally, in the memory retrieval phase, participants are asked to recall details from the witnessed event. In a pioneering experiment, Loftus (1975) tested 40 college students who watched a 3-min videotape depicting eight demonstrators who disrupted a class and then left the classroom. After the videotape, half of the participants received subtle misinformation by asking them a misleading question: "was the leader of four demonstrators a male?". The other half was asked: "was the leader of the twelve demonstrators a male?". One week later, all participants were interviewed about the number of demonstrators in the videotape. The second group falsely recalled on average two more demonstrators than the second group (average 8.85 compared with 6.4).

In the misinformation paradigm, false memories are caused by external misleading information and we term them *exogenous false memories*. These false memories have been found in all age populations (from infants to older people) through more than 40 years of investigation (Frenda, Nichols, & Loftus, 2011; Loftus, 2005). The misinformation paradigm focuses on false memories for details of an event. Yet, rich false memories of a wholly novel event can also be created using suggestive pressure. For instance, in the false memory implantation paradigm, participants have been presented with some fake evidence depicting a false event (e.g., a photo shopped old family picture) and then they are interviewed to elaborate on the false event. Otgaar, Candel, Merckelbach, and Wade (2009) presented children with a fake newspaper article about people being abducted by an UFO in their hometown when they were aged 4. The child participant was then told that his mother confirmed that he had been abducted by the UFO as well. Later, the participant was interviewed twice over seven days and asked to recall the UFO abduction. Many children (over 70%) vividly and falsely remembered that they were abducted by aliens (e.g., one child remembered seeing flashes, blue/green puppets, and other abducted children in the UFO).

Exogenous False Memories

Misinformation during Interview and Interrogation

External misleading information can be both verbal and nonverbal. During police interviews and interrogations, the phrasing of the questions as well as gestures made by the interviewers might undermine the accuracy of witnesses' memories. In one research line examining possible impacts from different types of questions (Kebbell, Evans, & Johnson, 2010; Kebbell & Johnson, 2000; Kebbell & Giles, 2000), participants first watched a short video of a crime such as a woman being attacked by a man; one week later, they answered "yes" or "no" to questions about the crime. Researchers found that negative questions (e.g., "Did the woman not have black hair?"), double negative questions (e.g., "Is it not true that the woman did not have black hair?") as well as leading questions in which the expected answer was given (e.g., "It is true to say that the attack happened in a park, isn't it?")

resulted in less accurate eyewitness reports compared with more simple questions (e.g., "Did the attack happen in a park?").

Sharman and Powell (2012) compared witnesses' susceptibility to misinformation across various phrasing structures of the interview questions. Participants went through the typical three-stage misinformation procedure (i.e., witnessing an event, receiving misinformation, and answering memory questions). Specifically, they were misinformed that there was an AJ's logo on the perpetrator's van when in fact there was an RJ's logo in the film. They constructed different types of questions containing misleading information. Of relevance here are the closed specific questions that require a "yes" or "no" response and contain specific misleading details at the same time (e.g., "Did Eric have an AJ's logo in large black letters on his van?") and open presumptive questions that suggest certain (misleading) information is true (e.g., "Tell me more about the AJ's logo on Eric's van"). They found that these two types of questions resulted in the highest false memory rates (38%) for the misinformation as well as the least accurate memories for correct details.

Nonverbal misinformation such as gestures during interviews can also lead to eyewitness memory distortions, which has recently been termed the *gestural misinformation effect* (Gurney, Pine, & Wiseman, 2013). In Gurney et al.'s (2013) study, participants watched footage of a crime scene and were later interviewed by an experimenter who acted as a police interviewer. During the interview, no verbal misinformation was given but when asked "Did you notice any jewellery?", the interviewer made either a gesture of a ring by pointing to a finger of the opposing hand or a gesture of a watch by grasping his wrist. The researchers found that more participants (30%) erroneously reported seeing a watch when a watch gesture was made than when a ring gesture was made (5%). Also, most participants (95%) reported seeing a ring when a ring gesture was made. In a similar research line, it was found that participants who saw a head nod of the interviewer reported higher confidence in their eyewitness reports than those who saw a head shake of the interviewer (Gurney, Vekaria, & Howlett, 2014).

More recently, Gurney, Ellis, and Vardon-Hynard (2016) examined whether subjective estimates of the nature and severity of the crime could be altered by misleading nonverbal information. Participants saw a video of a man punching another man in an alleyway and were then interviewed as eyewitnesses. The authors showed that a 'punching' gesture led participants to recall the crime more accurately, but a 'stabbing' gesture led more participants (61%) to recall that the victim was stabbed and severely injured compared with the punching condition (5.6%). The authors noted that gestural misinformation had the same and sometimes even a larger memory contaminating effect than verbal misinformation.

Misinformation concerning Eyewitness Identification

Misinformation can directly lead eyewitnesses to misidentify innocent people in a lineup. For example, Searcy, Bartlett, and Memon (2000) had participants look at a recording of an actual crime, the murder of an attendant at a dry cleaner's. Fifteen minutes later, participants had to listen to several narratives in which the witnessed crime was described. One narrative included misleading information that the perpetrator had a chipped tooth while in fact the perpetrator did not have a chipped tooth. Some hours later, participants were asked to identify the culprit in a lineup consisting of photographs of several suspects. Participants who received the misinformation were more likely to choose a person with a chipped tooth (25%) compared with those who did not receive the misinformation (6%).

Not only pre-identification misinformation (i.e., information provided before eyewitnesses make identifications from a lineup) undermines the accuracy of eyewitness memory, but also feedback after the eyewitness identification may distort eyewitness memory. In studies examining how post-identification feedback affects witnesses' memory reports (e.g., Erickson, Lampinen, Wooten, Wetmore, & Neuschatz, 2016; Skagerberg & Wright, 2009; Smalarz & Wells, 2014; Wells, Olson, & Charman, 2003), participants are provided with either confirming feedback (e.g., "Good, you identified the suspect") or no feedback after they identified a suspect from the lineup. The typical finding in these studies is that confirming feedback elevates participants' confidence in their memories and they are more willing to testify in court compared with the no feedback condition. Obviously, this confidence inflation can become a serious issue when the suspect is innocent.

Steblay, Wells, and Douglass (2014) conducted a meta-analysis of this post-identification effect based on data of 21 studies involving 7,000 participants from the United States, Canada, Europe, and Australia. They found that when an innocent person was chosen from a lineup, confirming feedback increased witnesses' memory clarity of the alleged culprit, memory for the alleged culprit's facial details, and their certainty in their (false) memories. The effect sizes of the post-identification effect on memory clarity and memory for facial details were medium to large in the reviewed studies (mean Cohen's *d* of 0.69 and 0.65, respectively).

Many studies on post-identification effects have been conducted in the artificial environment of the laboratory. However, Wright and Skagerberg (2007) tested whether eyewitnesses (victims and bystanders) to real crimes would change their responses to meta-memory questions after receiving feedback from the police. The authors evaluated actual eyewitnesses in the United Kingdom and observed that after police officers told the witnesses that they identified the true culprit, witnesses claimed better memories for faces and events compared with those who were told by the police that they did not identify the true culprit.

Misinformation from Co-witnesses

Crimes often involve multiple witnesses and hence, discussions among co-witnesses are common. In September 2003, a famous Swedish politician, Ann Lindh was murdered in a shopping mall. Witnesses discussed and influenced each other while they were kept in a room, such that the police collected erroneuous information about the identity of the perpetrator. The perpetrator was finally caught on the basis of DNA traces, but he did not match the descriptions of the witnesses.[4] Skagerberg and Wright (2008) studied the frequency of co-witness discussions at a UK Identification Suite. They found that 88% of the sampled eyewitnesses reported having seen co-witnesses at the crime scene and of these 58% discussed the crime with their co-witnesses on topics such as crime details and suspect details. This suggests that during such discussion, memory errors can easily be formed.

Indeed, discussions with co-witnesses can be a source of misinformation and thus influence witness' memory reports; a phenomenon that has been called memory conformity (for possible mechanisms, see Wright, Memon, Skagerberg, & Gabbert, 2009). Gabbert, Memon, and Allan (2003) first used a novel procedure where pairs of participants watched a different video of the same event; they were later encouraged to discuss the event with each other. The large majority (71%) of witnesses falsely recalled items acquired during the discussion with other co-witnesses. Witnesses who initiated the discussion were most likely to impact the other witness's memories (Gabbert, Memon, & Wright, 2006). Moreover,

[4] https://en.wikipedia.org/wiki/Memory_conformity

misinformation from familiar people (e.g., a friend or a romantic partner) has been shown to be more likely to be accepted than misinformation from a stranger (Hope, Ost, Gabbert, Healey, & Lenton, 2008). Recent research indicates that memory conformity is apparent in both children and adults (e.g., Otgaar, Howe, Brackmann, & van Helvoort, 2017).

Co-witness discussion can lead to eyewitness misidentification as well. Zajac and Henderson (2009) examined the impact of co-witness misinformation on lineup identification. Two witnesses viewed a video of a theft together and one witness (the confederate) falsely told the other that the thief had blue eyes when in fact the thief's eyes were brown. The authors found that witnesses who were misinformed by their co-witnesses were twice (47.2%) as likely to identify a blued-eye suspect as those who were not misinformed (23.6%). Eisen, Gabbert, Ying, and Williams (2017) had witnesses misinformed by co-witnesses that the perpetrator had a tattoo on his neck. They manipulated the retention interval between receiving the misinformation and lineup identification. They found that wrongful identifications of the tattooed person increased significantly at longer retention intervals. After a one-week delay, there were more witnesses who chose the innocent person with a tattoo (44%) than those who chose the true culprit (34%). Even when the co-witness seemed unreliable (e.g., consumed alcohol), witnesses still accepted their co-witness's misinformation and made wrongful identifications (Zajac, Dickson, Munn, & O'Neill, 2016).

Endogenous false memories

Apart from external misleading information, internal cognitive mechanisms can lead to the production of false memories. The typical research paradigm used to examine these endogenous false memories is the Deese/Roediger-McDermott (DRM) paradigm (Deese, 1959; Roediger & McDermott, 1995). In the DRM paradigm, participants are shown lists of associated words, such as *bed, rest, awake*, and later they are asked to recall/recognize which words were shown to them. Participants usually remember non-presented but related critical lure words such as *sleep* as the words they had seen with high confidence. Furthermore, they often falsely recollect these critical lures with rates that are indistinguishable from true memory rates (Roediger & McDermott, 1995). The false memory effect in the DRM paradigm has been shown to be a robust phenomenon in children and adults (Howe, 2005, 2006), using different stimuli (Hege & Dodson, 2004; Schacter, Israel, & Racine, 1999).

We term this type of memory illusions “endogenous” as the theoretical idea is that these illusions are caused by automatic spreading activation of mental representations (Howe, Wimmer, Gagnon, & Plumpton, 2009; Roediger, Balota, & Watson, 2001). That is, when witnesses view some items, related but not presented concepts will be automatically activated and this might create false memories of non-presented items. For example, Otgaar, Howe, Brackmann, and Smeets (2016) showed participants a video about a robbery in which a culprit entered the cafeteria and demanded money from the people at the cash desk. Associated items such as money, cahier, black jacket, masked hat, and a robber were shown in the video. However, without any misinformation, participants automatically formed a false memory for the presence of a gun in the video.

Emotion and False Memory

Emotion is one important factor that drives endogenous false memories. From a forensic perspective, this issue is relevant as people generally experience intense and/or negative emotions when facing a crime. Research found that 90% of the participants formed false memories for negative public events (e.g., the 911 terrorist attack), but only 41.7% of the participants had false memories for positive public events (Porter, Taylor, & ten Brinke, 2008). Research examining the effect of emotion on the production of spontaneous false memories present participants with different emotionally-laden lists (negative, positive) and then examine participants’ susceptibility in forming false memories. In general, research has revealed that false recognition rates for negative DRM lists are higher than for positive or neutral DRM lists (Brainerd, Holliday, Reyna, Yang, & Toglia, 2010; Brainerd, Stein, Silveira, Rohenkohl, & Reyna, 2008; Howe, Candel, Otgaar, Malone, & Wimmer, 2010).

A crime scene may elicit not only a negatively valenced emotion such as fear and anger, but often also will induce high arousal. Brainerd et al. (2010) manipulated both valence and arousal of DRM lists. They found that negative emotion generated higher false memory rates than positive emotion and high arousal generated higher false memory rates than low arousal. Bookbinder and Brainerd (2017) administered negative, neutral, and positive pictures to participants while controlling the arousal level of the pictures. Negative pictures, like negative words, enhanced false memory on both an immediate and a one-week delay recognition tests. On the basis of the studies summarized in this section, we can safely conclude that both negative valence and high arousal enhance the production of false memories (Bookbinder & Brainerd, 2016; Kaplan, Van Damme, Levine, & Loftus, 2016).

Stress and False Memory

As negative emotional material fuels false memory formation, one might expect that stress – which is often experienced as negative – promotes false memory as well. However, studies examining the effects of stress on false memory have found mixed results. Payne, Nadel, Allen, Thomas, and Jacobs (2002) were the first to examine the effect of stress on false memory creation. In their study, participants were asked to give a speech so as to induce moderate psycho-social stress. Later, participants listened to DRM lists and then received a recognition test. Stress increased false memory rates compared with a no-stress condition.

However, this pattern has not been replicated in other studies. For example, Smeets, Jelicic, and Merckelbach (2006) followed a similar procedure as in Payne et al.'s (2002) study – a stress induction phase, a DRM study phase and a memory test phase. They also collected participants' cortisol levels, which is a biological indicator of stress, at several times during the experiment as a check on the stress induction manipulation. In two studies, the authors did not find any evidence that stress increased false memory production. Furthermore, Smeets, Otgaar, Candel, and Wolf (2008) exposed participants to the cold pressor stress task (CPS) in which participants have to immerse their arm in ice-cold water for as long as possible. Again, there was no indication that false memory proneness was affected by levels of stress.

It seems that stress does not increase endogenous false memories, but it might impair true memories for peripheral details so that it makes witnesses highly susceptible to misinformation (i.e., creating exogenous false memories) (Kaplan et al., 2016). Morgan, Southwick, Steffian, Hazlett, and Loftus (2013) examined over 800 military personnel's false memories for highly stressful events. Participants went through a highly stressful interrogation where they acted as a mock prisoner of war and were treated with physical assaults. Following the stressful event, a misinformation questionnaire was introduced and later participants' memories for the aggressive interrogator were assessed. Around a half of the participants who received the misinformation identified the wrong individual as their interrogator.

Prevention and Identification of False memories

Preventing False Memory and Promoting Accurate Memory

The story so far is that false memories can be easily generated. However, researchers have devised several ways to prevent the occurrence of false memories and promote the retrieval of accurate memories. A general principle is to avoid giving suggestive information to witnesses during investigative interviews. One important step here is the construction of empirically validated interview protocols that maximize accurate reporting and minimize false reports. One well-studied interview protocol is the Cognitive Interview (CI). The CI is a protocol for interviewing witnesses and has been studied for more than 30 years. The CI is composed of several cognitive principles that may enhance accurate statements. During the CI, eyewitnesses undergo the following procedure (for details see Fisher & Schreiber, 2007). First, the interview starts in a friendly manner to build rapport with the witness, which will lower the stress that the witness experiences when facing a police investigator. Research has demonstrated that rapport building during CI decreases a witness' susceptibility to misinformation for a mock-crime (Vallano & Compo, 2011). Next, the witness is encouraged to report everything recalled, without interruption from the interviewer, so the witness controls the flow of information instead of being led by the interviewer. Following this free-narrative phase, the interviewer probes the witness about the target event with open-ended questions, which, as reviewed above, leads to fewer false memories than closed questions. Memon, Meissner, and Fraser (2010) reviewed 25 years' laboratory and field studies on the CI, and found that the CI led to a large and significant increase in correct details with only a small increase in errors compared with standard interviewing conditions.

Second, post-warnings are found to be effective in reducing false memories caused by misinformation. Post-warnings refer to warnings given to participants to the effect that some of the post-event information they received might be inaccurate. For instance, participants who received misinformation from their co-witnesses were warned later that their co-witnesses might have watched a different video, thus making the participants to reflect on their own memories (Paterson, Kemp, & Mclntyre, 2012). Blank and Launay (2014) conducted a meta-analysis of 25 studies from 1980s to 2010s on the effect of post-warnings. They found that post-warnings can reduce the original memory misinformation effect to 43% of its original (no-warning) size.

Third, using a blind lineup administration can prevent witnesses' memories from distortion during lineup identification. In a blind lineup administration, the administrator does not know the identity of the suspect. A blind lineup can prevent the administrator from giving subtle hints such as an unconscious gesture to the witness. Thus, in a blind procedure, it is less likely that the administrator will intentionally or unintentionally lead the witness to identify a person on the basis of misinformation than during a non-blind lineup. Blind lineup administration can also reduce the post-identification effect such that witnesses' confidence and judgements about their identifications do not escalate due to erroneous feedback (Dysart, Lawson, & Rainey, 2012).

Distinguishing between True versus False memories

False memories have been reported to contain fewer sensory details than true memories (e.g., Norman & Schacter, 1997), but there are also many cases where false memories are experienced as vividly like true memories (Foley, Bays, Foy, & Woodfield, 2015). With the development of brain scanning techniques such as functional magnetic resonance imaging (fMRI), the neural signatures of true and false memories provide possibilities to differentiate false from true memories. Neural correlates of true and false memories have been studied extensively in recent years. Slotnick and Schacter (2004; 2006) identified different activations in the sensory-processing brain areas for true and false memories. Similar to the DRM paradigm, participants in their studies viewed various shapes in the study phase, and then formed false memories for related but not presented shapes in the test phase. fMRI scanning of the test phase revealed that true memories had greater activation in early visual processing regions (Brodmann area 17, 18) than false memories. Okado, Stark, and Loftus (2010) used the misinformation paradigm where they presented participants with picture stimuli in the study phase, followed by misinformation one day later. They also found that true memories of visual stimuli were preferentially associated with early visual processing areas, which are normally involved in sensory encoding of visual stimuli (see also Atkins & Reuter-Lorenz, 2011).

Other studies showed that true memories for auditory stimuli were associated with activation in auditory sensory processing regions (e.g., left temporo-parietal cortex; Cabeza et al., 2001; Abe, Okuda, Suzuki, et al, 2008). On the basis of this type of results, Schacter, Chamberlain, Gaesser, and Gerlach (2011) proposed the *sensory reactivation hypothesis*, which holds that true memories are accompanied by retrieval of more sensory/perceptual details than false memories. This pattern would manifest itself in the reactivation of sensory/perceptual encoding brain regions that were engaged during the establishment of true but not false memories. Thus, when people have truly seen or heard target stimuli, brain areas that were engaged in processing the stimuli (e.g., early visual cortex) will be activated as soon as they try to retrieve memories of the targets. False memories lack such kind of activations as they have not been “seen” or “heard” before. The sensory reactivation hypothesis has been supported by recent studies (Dennis, Bowman, & Vandekar, 2012; Dennis, Johnson, & Peterson, 2014).

Researchers have explored the unique neural signature that is associated with false memories as well. In a recent study, Chadwick, Anjum, Kumaran, Schacter, Spiers, and Hassabis (2016) used fMRI to search for a neural code for false memories in the DRM paradigm. They manipulated the semantic overlap between studied items and critical lures from low to high. The computational analysis enabled them to test the neural overlap between DRM items and critical lures, corresponding to the semantic overlap between them. They found that patterns of activity in the temporal pole can predict false memories and that subject-specific temporal pole neural coding can predict individual false memories.

However, researchers are cautious when it comes to applying neuroimaging techniques in the courtroom to identify an individual’s memory as true or false. First, neuroimaging studies in the lab normally examine true and false memories for simple stimuli such as words and pictures, and brain activations induced by simple stimuli might be very different from activations of rich events such as a crime (Schacter & Loftus, 2013). What’s more, although researchers found neural differences between true and false memories, those differences are based on the summaries of brain activities in a group of participants, making it hard to apply the results to a single participant (Van de Ven, Otgaar, & Howe, 2018). Recently, there are studies showing neural decoding of individual (false) memories (e.g., Chadwick et al., 2016), but the differentiation between false and true memories is at present far from 100% accurate. Still, as neuroimaging techniques develop and more complex stimuli are examined, it looks promising to distinguish false from true memories at the neural level, particularly because it is almost impossible to distinguish false from true memories at the behavioral level (Bernstein & Loftus, 2009).

Conclusions and Future Directions for China

We reviewed two types of false memories (exogenous vs. endogenous) and their possible consequences in eyewitness testimony. Exogenous false memories may occur after people receive external misinformation, which can be suggestive questions or gestures during interview, misleading information pre- or post-lineup identification, or false information from a co-witness. Endogenous false memories are generated by internal cognitive mechanisms without external misleading information, and can be inflated by negative emotion and high arousal. Overall, false memories can lead to wrong descriptions of the perpetrator or the crime that may guide the investigation into a wrong direction, or more directly, cause eyewitnesses to misidentify an innocent person as the perpetrator. We also reviewed that designated interview protocols such as the CI, blind lineup administration, and post-warnings could prevent or lower the chance of false memory occurrence.

In returning now to the cases that were presented at the beginning, we may find several factors relevant to what we have reviewed here. For example, Haisheng Zhang was not only misidentified by the victim, but also by three other teenage girls who lived in the same village. Being co-witnesses who knew each other well, the girls probably talked to each other and reached memory conformity in the end. Moreover, Zhang's lawyer presented evidence at the court that the police said to the girls to "look carefully at the shoes" during the lineup identification, which might be a suggestive hint, but the court seems to have overlooked this and convicted Zhang as guilty of rape.

In the case of Jibin Xu, the court relied heavily on the statements of the victim witness that Xu was the perpetrator. At Xu's first trial, he proposed that the victim was lying to imprison him, yet the court was not convinced by this alternative explanation. It is unknown whether the witness was lying or merely had a false memory, but this case is very similar to the many cases archived in the Innocence Project (www.innocenceproject.org) where witnesses had false memories about the perpetrators. If the risk of false memory had been widely acknowledged by practitioners in the legal system at that time, Xu might have avoided the destiny of spending 8 years in prison.

Fortunately, lessons have already been learnt by countries in North America and Europe, and those lessons might inform China about what are the best to protect innocent people from being criminalized by false memories. For instance, in the United States, the supreme court of New Jersey issued a ruling that the unreliability of memory should be taken into account when evaluating eyewitness identification evidence in court (State v. Henderson, 2011). The Criminal Procedure Law of the P.R. China has been revised in 2012, in which eyewitness testimony is listed as one of the main categories of evidence (Article 42) and the testimony of a witness may be used as a basis in deciding a case under certain circumstances (Article 47), but no specific rule is written in regulating eyewitness identification processes such as lineup administration. In practice, the Public Security Organs and the People's Procuratorate provide provisions that the identification should be hosted by investigation or prosecution personnel, and 97% of the identifications in China are carried out by investigators who undertook the case (Chen, 2015). The aforementioned situations may be prone to the risk factors of false memories such as unintentional misinformation and suggestion. In our opinion, several steps are needed to increase awareness concerning the importance of eyewitness testimony and false memories in Chinese legal cases.

We contend that improving awareness of false memories in the legal arena is perhaps the first step to start. Memories are more prone to errors than many people think. It is especially important for judges, lawyers, and the police to be *aware* of that. Knowledge about how memory works and how to prevent false memories can be shared in ways of workshops and seminars (Loftus, 2003). This is important as many legal professionals possess flawed ideas concerning the functioning of memory (i.e., memory is like video-taping). A first direction could be to educate legal professionals such as the police about the science of memory and its relevance to courts of law. Such interventions might help legal professionals to get rid of their biases regarding the functioning of memory (Lilienfeld, Ammirati, & Landfield, 2009). A second important follow-up step would be to collaborate closely with legal professionals and attempt to launch several actions in investigative and juridical processes to *prevent* false memories, such as the use of empirically-validated interview protocols and blind lineup identification and launching new laws and regulations on organization and administration of eyewitness identification. Such actions have already taken place in countries such as the UK and the Netherlands.

Finally, other measures can be taken by actively *recognizing* the possibility of false memory occurrence in legal practice. An ideal route to accomplish this is when triers of fact more often consult memory experts in legal cases. In many countries, expert witnesses who are memory researchers as well are called upon to provide their expert opinion concerning a memory-related issue in a case such as the disclosure of child's statements on sexual abuse (Otgaar & Howe, 2018). Such experts might considerably assist judges and lawyers and might help judges in reaching legal decisions grounded in memory science. For instance, Wise and Safer (2012) designed a toolkit to analyze the trustworthiness of eyewitness testimony by evaluating the risky factors step by step that we have reviewed above. Criminal proceedings are to a large extent dependent on what eyewitnesses report. The purpose of the review is not to leave the impression that eyewitnesses are wrong all the time or even most of the times. Eyewitnesses may often attain impressive accuracy and in many cases, eyewitnesses contribute critically to fair and just legal proceedings. Our review of the literature was an attempt to further increase the trust that triers of fact can place in eyewitnesses by excluding conditions that promote false memories.

References

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

Atkins, A. S., & Reuter-Lorenz, P. A. (2011). Neural mechanisms of semantic interference and false recognition in short-term memory. *NeuroImage*, *56*, 1726-1734.

Bernstein, D. M., & Loftus, E. F. (2009). How to tell if a particular memory is true or false. *Perspectives on Psychological Science*, *4*, 370-374.

Blank, H., & Launay, C. (2014). How to protect eyewitness memory against the misinformation effect: A meta-analysis of post-warning studies. *Journal of Applied Research in Memory and Cognition*, *3*, 77-88.

Bookbinder, S. H., & Brainerd, C. J. (2016). Emotion and false memory: The context–content paradox. *Psychological Bulletin*, *142*, 1315-1351.

Bookbinder, S. H., & Brainerd, C. J. (2017). Emotionally negative pictures enhance gist memory. *Emotion*, *17*, 102-119.

Brainerd, C. J., Holliday, R. E., Reyna, V. F., Yang, Y., & Toglia, M. P. (2010). Developmental reversals in false memory: Effects of emotional valence and arousal. *Journal of Experimental Child Psychology*, *107*, 137-154.

Brainerd, C. J., Stein, L. M., Silveira, R. A., Rohenkohl, G., & Reyna, V. F. (2008). How does negative emotion cause false memories? *Psychological Science*, *19*, 919-925.

Cabeza, R., Rao, S. M., Wagner, A. D., Mayer, A. R., & Schacter, D. L. (2001). Can medial temporal lobe regions distinguish true from false? An event-related functional MRI study of veridical and illusory recognition memory. *Proceedings of the National Academy of Sciences*, *98*, 4805-4810.

Chadwick, M. J., Anjum, R. S., Kumaran, D., Schacter, D. L., Spiers, H. J., & Hassabis, D. (2016). Semantic representations in the temporal pole predict false memories. *Proceedings of the National Academy of Sciences*, *113*, 10180-10185.

Chen, X. (2015). *Eyewitness identification research: from the perspectives of Law and Psychology* (*pp.*83–84). Beijing, China: China Procuratorate Press. (in Chinese)

Chen, Y. (2007). A perspective on miscarriage of justice in China: An analysis based on 20 famous wrongful convicted cases. *China Legal Science*, *3*, 45-61. (in Chinese)

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*, 17–22.

Dennis, N. A., Bowman, C. R., & Vandekar, S. N. (2012). True and phantom recollection: an fMRI investigation of similar and distinct neural correlates and connectivity. *Neuroimage*, *59*, 2982-2993.

Dennis, N. A., Johnson, C. E., & Peterson, K. M. (2014). Neural correlates underlying true and false associative memories. *Brain and Cognition*, *88*, 65-72.

Dysart, J. E., Lawson, V. Z., & Rainey, A. (2012). Blind lineup administration as a prophylactic against the postidentification feedback effect. *Law and Human Behavior*, *36*, 312-319.

Eisen, M. L., Gabbert, F., Ying, R., & Williams, J. (2017). "I think he had a tattoo on his neck": How co-witness discussions about a perpetrator's description can affect eyewitness identification decisions. *Journal of Applied Research in Memory and Cognition*, *6*, 274-283

Erickson, W. B., Lampinen, J. M., Wooten, A., Wetmore, S., & Neuschatz, J. (2016). When snitches corroborate: effects of post-identification feedback from a potentially compromised source. *Psychiatry, Psychology and Law*, *23*, 148-160.

Fisher, R. P., & Schreiber, N. (2007). Interviewing protocols to improve eyewitness memory. *The handbook of eyewitness psychology*, *1*, 53-80.

Foley, M. A., Bays, R. B., Foy, J., & Woodfield, M. (2015). Source misattributions and false recognition errors: Examining the role of perceptual resemblance and imagery generation processes. *Memory*, *23*, 714-735.

Frenda, S. J., Nichols, R. M., & Loftus, E. F. (2011). Current issues and advances in misinformation research. *Current Directions in Psychological Science*, *20*, 20-23.

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.

Gabbert, F., Memon, A., & Wright, D. B. (2006). Memory conformity: Disentangling the steps toward influence during a discussion. *Psychonomic Bulletin & Review*, *13*, 480-485.

Garrett, B. L. (2011). *Convicting the innocent*. Cambridge, MA: Harvard University Press.

Gurney, D. J., Ellis, L. R., & Vardon-Hynard, E. (2016). The saliency of gestural misinformation in the perception of a violent crime. *Psychology, Crime & Law*, *22*, 651-665.

Gurney, D. J., Pine, K. J., & Wiseman, R. (2013). The gestural misinformation effect: skewing eyewitness testimony through gesture. *The American Journal of Psychology*, *126*, 301-314.

Gurney, D. J., Vekaria, K. N., & Howlett, N. (2014). A nod in the wrong direction: Does non-verbal feedback affect eyewitness confidence in interviews *Psychiatry, Psychology and Law*, *21*, 241-250.

Hege, A. C., & Dodson, C. S. (2004). Why distinctive information reduces false memories: evidence for both impoverished relational-encoding and distinctiveness heuristic accounts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 787-795.

Hope, L., Ost, J., Gabbert, F., Healey, S., & Lenton, E. (2008). "With a little help from my friends…": The role of co-witness relationship in susceptibility to misinformation. *Acta Psychologica*, *127*, 476-484.

Howe, M. L. (2005). Children (but not adults) can inhibit false memories. *Psychological Science*, *16*, 927-931.

Howe, M. L. (2006). Developmentally invariant dissociations in children's true and false memories: Not all relatedness is created equal. *Child Development*, *77*, 1112-1123.

Howe, M. L., Candel, I., Otgaar, H., Malone, C., & Wimmer, M. C. (2010). Valence and the development of immediate and long-term false memory illusions. *Memory*, *18*, 58-75.

Howe, M. L., & Knott, L. (2015). The fallibility of memory in judicial processes: Lessons from the past and their modern consequences. *Memory, 23*, 633–656.

Howe, M. L., Knott, L. M., & Conway, M. A. (2018). *Memory and miscarriages of justice.* Abingdon, UK: Routledge.

Howe, M. L., Wimmer, M. C., Gagnon, N., & Plumpton, S. (2009). An associative-activation theory of children's and adults' memory illusions. *Journal of Memory and Language*, *60*, 229-251.

Kaplan, R. L., Van Damme, I., Levine, L. J., & Loftus, E. F. (2016). Emotion and false memory. *Emotion Review*, *8*, 8-13.

Kebbell, M. R., & Giles, D. C. (2000). Some experimental influences of lawyers' complicated questions on eyewitness confidence and accuracy. *The Journal of Psychology, 134*, 129-139.

Kebbell, M. R., Evans, L., & Johnson, S. D. (2010). The influence of lawyers' questions on witness accuracy, confidence, and reaction times and on mock jurors' interpretation of witness accuracy. *Journal of Investigative Psychology and Offender Profiling*, *7*, 262-272.

Kebbell, M. R., & Johnson, S. D. (2000). Lawyers' questioning: The effect of confusing questions on witness confidence and accuracy. *Law and Human Behavior*, *24*, 629-641.

Lilienfeld, S. O., Ammirati, R., & Landfield, K. (2009). Giving debiasing away: Can psychological research on correcting cognitive errors promote human welfare? *Perspectives on Psychological Science*, *4*, 390-398.

Loftus, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology*, *7*, 560-572.

Loftus, E. (2003). Our changeable memories: Legal and practical implications. *Nature Reviews. Neuroscience*, *4*, 231-234.

Loftus, E. F. (2004). Memories of things unseen. *Current Directions in Psychological Science*, *13*, 145-147.

Loftus, E. F. (2013). 25 years of eyewitness science…… finally pays off. *Perspectives on Psychological Science*, *8*, 556-557.

Memon, A., Meissner, C. A., & Fraser, J. (2010). The Cognitive Interview: A meta-analytic review and study space analysis of the past 25 years. *Psychology, Public Policy, and Law, 16*, 340-372.

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.

Norman, K. A., & Schacter, D. L. (1997). False recognition in younger and older adults: Exploring the characteristics of illusory memories. *Memory & Cognition*, *25*, 838-848.

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.

Otgaar, H., & Howe, M.L. (2018). *Finding the truth in the courtroom: Dealing with deception, lies, and memories*. New York: Oxford University Press.

Otgaar, H., Howe, M. L., Brackmann, N., & Smeets, T. (2016). The malleability of developmental trends in neutral and negative memory illusions. *Journal of Experimental Psychology: General, 145*, 31.

Otgaar, H., Howe, M. L., Brackmann, N., & van Helvoort, D. H. (2017). Eliminating age differences in children's and adults' suggestibility and memory conformity effects. *Developmental Psychology*, *53*, 962-970.

Otgaar, H., Sauerland, M., & Petrila, J. P. (2013). Novel shifts in memory research and their impact on the legal process: introduction to the special issue on memory formation and suggestibility in the legal process. *Behavioral Sciences & the Law, 31*, 531-540.

Paterson, H. M., Kemp, R., & McIntyre, S. (2012). Can a witness report hearsay evidence unintentionally? The effects of discussion on eyewitness memory. *Psychology, Crime & Law*, *18*, 505-527.

Peterson, J. L., Hickman, M. J., Strom, K. J., & Johnson, D. J. (2013). Effect of forensic evidence on criminal justice case processing. *Journal of Forensic Sciences, 58*, 78-90.

Payne, J. D., Nadel, L., Allen, J. J., Thomas, K. G., & Jacobs, W. J. (2002). The effects of experimentally induced stress on false recognition. *Memory*, *10*, 1-6.

Porter, S., Taylor, K., & ten Brinke, L. (2008). Memory for media: Investigation of false memories for negatively and positively charged public events. *Memory, 16*, 658-666.

Roediger III, H. L., Balota, D. A., & Watson, J. M. (2001). Spreading activation and arousal of false memories. In H. L. Roediger III, J. S. Nairne, I. Neath, & A. M. Surprenant (Eds), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 95-115). Washington DC: American Psychological Association.

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*, 803–14.

Schacter, D. L. (2012). Adaptive constructive processes and the future of memory. *American Psychologist*, *67*, 603–613.

Schacter, D.L., Chamberlain, J., Gaesser, B. & Gerlach, K. Neuroimaging of true, false, and imaginary memories. (2012). In L. Nadel & W.P. Sinnott-Armstrong(Eds), *Memory and Law* (pp. 233-262). Oxford University Press, New York.

Schacter, D. L., Israel, L., & Racine, C. (1999). Suppressing false recognition in younger and older adults: The distinctiveness heuristic. *Journal of Memory and Language, 40*, 1-24.

Schacter, D. L., & Loftus, E. F. (2013). Memory and law: what can cognitive neuroscience contribute? *Nature Neuroscience*, *16*, 119-123.

Searcy, J., Bartlett, J. C., & Memon, A. (2000). Influence of post-event narratives, line-up conditions and individual differences on false identification by young and older eyewitnesses. *Legal and Criminological Psychology*, *5*, 219-235.

Sharman, S. J., & Powell, M. B. (2012). A comparison of adult witnesses' suggestibility across various types of leading questions. *Applied Cognitive Psychology*, *26*, 48-53.

Skagerberg, E. M., & Wright, D. B. (2008). The prevalence of co-witnesses and co-witness discussions in real eyewitnesses. *Psychology, Crime & Law*, *14*, 513-521.

Skagerberg, E. M., & Wright, D. B. (2009). Susceptibility to postidentification feedback is affected by source credibility. *Applied Cognitive Psychology*, *23*, 506-523.

Slotnick, S. D., & Schacter, D. L. (2004). A sensory signature that distinguishes true from false memories. *Nature Neuroscience*, *7*, 664-672.

Slotnick, S. D., & Schacter, D. L. (2006). The nature of memory related activity in early visual areas. *Neuropsychologia*, *44*, 2874-2886.

Smalarz, L., & Wells, G. L. (2014). Post-identification feedback to eyewitnesses impairs evaluators’ abilities to discriminate between accurate and mistaken testimony. *Law and Human Behavior*, *38*, 194-202.

Smeets, T., Jelicic, M., & Merckelbach, H. (2006). The effect of acute stress on memory depends on word valence. *International Journal of Psychophysiology*, *62*, 30-37.

Smeets, T., Otgaar, H., Candel, I., & Wolf, O. T. (2008). True or false? Memory is differentially affected by stress-induced cortisol elevations and sympathetic activity at consolidation and retrieval. *Psychoneuroendocrinology*, *33*, 1378-1386.

Stark, C. E., Okado, Y., & Loftus, E. F. (2010). Imaging the reconstruction of true and false memories using sensory reactivation and the misinformation paradigms. *Learning & Memory*, *17*, 485-488.

State v. Henderson (2011), 208 N.J. 208.

Steblay, N. K., Wells, G. L., & Douglass, A. B. (2014). The eyewitness post identification feedback effect 15 years later: Theoretical and policy implications. *Psychology, Public Policy, and Law*, *20*, 1-18.

Vallano, J. P., & Compo, N. S. (2011). A comfortable witness is a good witness: Rapport-building and susceptibility to misinformation in an investigative mock-crime interview. *Applied Cognitive Psychology*, *25*, 960-970.

Van de Ven, V., Otgaar, H., & Howe, M. L. (2018). A neurobiological account of false memories. In H. Otgaar & M. L. Howe (Eds.), *Finding the truth in the courtroom: Dealing with deception, lies, and memories* (pp. 75-99). New York: Oxford University Press.

Wells, G. L., Olson, E. A., & Charman, S. D. (2003). Distorted retrospective eyewitness reports as functions of feedback and delay. *Journal of Experimental Psychology: Applied*, *9*, 42-52.

Wise, R. A., & Safer, M. A. (2012). A method for analyzing the accuracy of eyewitness testimony in criminal cases. *Court Review*, *48*, 22-34.

Wright, D. B., & Skagerberg, E. M. (2007). Postidentification feedback affects real eyewitnesses. *Psychological Science*, *18*, 172-178.

Wright, D. B., Memon, A., Skagerberg, E. M., & Gabbert, F. (2009). When eyewitnesses talk. *Current Directions in Psychological Science*, *18*, 174-178.

Zajac, R., Dickson, J., Munn, R., & O'neill, S. (2016). Trussht me, I know what I sshaw: The acceptance of misinformation from an apparently unreliable co-witness. *Legal and Criminological Psychology*, *21*, 127-140.

Zajac, R., & Henderson, N. (2009). Don't it make my brown eyes blue: Co-witness misinformation about a target's appearance can impair target-absent line-up performance. *Memory*, *17*, 266-278.

CHAPTER 3

Positive Consequences of True and False Memories

This Chapter is an adapted version of the following paper:

Wang, J., Otgaar, H., Howe, M. L., Lippe, F., & Smeets, T. (2018). The nature and consequences of false memories for visual stimuli. *Journal of Memory and Language*, *101,* 124-135.

Abstract

Memories can prime cognitive tasks such as the perceptual closure task. Different theoretical views exist regarding whether false memories contain perceptual information or are merely conceptual in nature. To address this question, we conducted three experiments to examine whether false memories for pictures had a priming effect on a perceptual closure task. In Experiment 1, participants were presented with pictorial versions of Deese/Roediger-McDermott (DRM) lists and received a recognition task. Finally, in the perceptual closure task (PCT), participants were shown degraded pictures (studied pictures, critical pictures, unrelated pictures) that became clearer over time and had to identify the object depicted as quickly as possible. The results showed that false memories for pictures did not exhibit a priming effect in the PCT. Specifically, picture identifications based on false memories for visual stimuli were significantly slower than those based on true memories and the former did not differ from that of unrelated items. In Experiments 2 and 3, we manipulated the modality (verbal vs. pictorial) of the study phase and the PCT phase. In both experiments, false memories for pictures primed pictures significantly slower than true memories in the pictorial PCT, but false memories for pictures primed words faster than true memories in the verbal PCT. Our results suggest that false memories for pictures are unlikely to contain perceptual information but rather that they are conceptual in nature.

Keywords: false memory, memory representation, perceptual closure task, positive consequences

It is not uncommon that you falsely remember when you had your meeting or where you dated your girlfriend a year ago. What matters is whether these memory distortions will result in any unwanted consequences. False memories are commonly associated with negative consequences such as false accusations of child sexual abuse (Otgaar, Sauerland, & Petrila, 2013). However, in recent years, studies have accumulated showing that false memories have adaptive value and can be beneficial under certain circumstances (Howe, 2011; Schacter, 2012). As many of these beneficial effects of false memories have been observed with verbal stimuli, the general aim of the current study is to investigate whether these same positive outcomes generalize to *visual* stimuli.

For example, a series of studies have demonstrated that false memories facilitate problem-solving behavior equivalent to or better than true memories (Howe, Garner, Dewhurst, & Ball, 2010; Howe, Garner, & Patel, 2013; Howe, Wilkinson, Garner, & Ball, 2016; Wang, Otgaar, Howe, Smeets, Merkelbach, & Nahouli, 2016). In these studies, participants are usually presented with associated lists (e.g., *women, husband, uncle, lady, male*) using the well-known Deese/Roediger-McDermott paradigm (DRM; Deese, 1959; Roediger & McDermott, 1995). After the presentation of such DRM lists, participants commonly falsely recollect the non-presented ‘critical lure’ word (i.e., *man*) which is associatively related to each of the list words. Then, participants are asked to solve compound remote associate task problems. Each problem is a three-word puzzle (e.g., *old/hole/super*) and they have to come up with a theme word (in this case, *man*) that could link all the three words. Importantly, when the solutions are the non-presented critical lures, participants solve these problems in a more efficient way than when the solutions are the words presented on the lists. Not only when using compound remote associate problems do false memories exert positive consequences, but also with analogical reasoning problems in both children (Howe, Threadgold, Norbury, Garner, & Ball, 2013) and adults (Howe, Garner, Threadgold, & Ball, 2015).

Additional studies have concentrated on the positive priming effects of false memories on an adapted perceptual closure task (PCT) – a measure loosely linked to intelligence (Otgaar, Howe, Beers, van Hoof, Bronzwaer, & Smeets, 2015). In this PCT, participants are presented with degraded stimuli that become clearer over time and they have to indicate what the stimuli are (e.g., a certain word) as soon as they recognize the stimuli. The PCT requires people to fill in the missing parts of a degraded image, which is similar to subtasks in certain intelligence tasks (Luteijn & Barelds, 2004). In Otgaar et al.'s (2015) study, participants received DRM lists and a recognition test and they were then exposed to the PCT (Experiment 1). Here, they were presented with degraded versions of both presented words and the critical lures. The significant finding was that false memories for the critical lures primed the PCT faster than the presented words. The same finding emerged when no recognition task was included (Experiment 2). Again, this is evidence indicating the processing advantages associated with false over true memories.

The above lines of research clearly demonstrate the positive value of false memories. However, these findings may be somewhat limited inasmuch as the stimuli used in these studies were purely conceptual (i.e., words) and no studies have examined whether false memories for pictures have any positive consequence on cognitive tasks such as the PCT. This question is important for the following reasons. First, research dealing with visual false memories is limited (see Miller & Gazzaniga, 1998), but studies that have been conducted reveal that false memory rates for pictures differ from false memories for words. For example, Israel and Schacter (1997) found that false memories for pictures resulted in significantly lower acceptance rates than verbal false memories (see also Howe, 2006). When false memories are derived from visual scenes, recognition rates for critical lures were relative low as well (~30%; Otgaar, Howe, Peters, Smeets, & Moritz, 2014). One possible explanation for these differences in false memory rates is distinctiveness. That is, pictures contain more details which are distinct from each another and this tends to weaken the associatively-based generation of false memories (Koutstaal, Schacter, & Brenner, 2001; Seamon, Luo, Schlegel, Greene, & Goldenberg, 2000). So, the cognitive processes underlying visual and verbal false memories might be different, and hence, it is unknown whether false memories for visual stimuli also exert positive influences on subsequent priming tasks.

Second, spreading activation theories (Howe, Wimmer, Gagnon, & Plumpton, 2009; Roediger, Balota, & Watson, 2001) have been put forward to explain the positive consequences of associatively generated (e.g., verbal) false memories. In these activation theories, DRM list members (nodes) are embedded within an associative network. Activation of one conceptual node spreads to other related nodes. Activation of related but non-presented concepts not only leads to false recall/recognition of non-presented words but also triggers additional spreading activation that impacts performance on related (e.g., associative) tasks (Howe et al., 2010; Otgaar et al., 2015). When visual stimuli are presented and visual false memories are induced, it is unclear whether visual representations are integrated in this network and thus, it is unclear whether they will result in any positive priming consequences.

In theories about how information is represented in memory (e.g., Dual Coding Theory), it is thought that pictures are encoded in parallel into both imaginal and verbal representations (Paivio, 1991; Sadoski & Paivio, 2004). If false memories derived from visual stimuli also contain perceptual representations, false memories for pictures should facilitate priming on imagery/perceptual tasks. Studying the consequences of visual false memories on a priming task might help uncover the mechanism underlying false memories.

Perceptual vs. Conceptual False Memories

Indeed, there is no consensus regarding whether false memories encompass a perceptual mental representation in the same way that true memories do. One line of research favors the view that the memory representation for DRM critical lures contains perceptual details. For instance, McDermott (1997) had participants study DRM lists and then perform a word-stem completion task (e.g., bas___ for the target *basket*) or a fragment completion task (e.g., b_s___ for the target *basket*). These tasks were referred to as perceptual priming tasks by the author. The results showed that critical lures had a significant priming effect on the stem/fragment completion rates as compared to non-studied items, although the level of priming was not as strong as studied items. McKone and Murphy (2000) replicated McDermott’s (1997) results and even found that the magnitude of critical lures’ priming effect was equivalent to that of studied items. The priming effects of critical lures were also found in other studies involving similar tasks such as a lexical decision task (i.e., to classify stimuli as words or nonwords) (Tajika, Neumann, Hamajima, & Iwahara, 2005; Tse & Neely, 2005, 2007).

The opponent view is that false memories for critical lures do not contain perceptual information. For example, Hicks and Starns (2005) manipulated the study modality of DRM lists (visual vs. auditory presentation) to examine whether critical lure priming was modality specific. The logic was that, if critical lures are perceptually encoded, visual DRM lists should induce visually encoded critical lures and auditory lists should lead to auditory encodings of critical lures. Thus, critical lure priming of the *visual* stem completion task should only exist in the visual study modality but not in the auditory study modality. Hicks and Starns (2005) found that critical lures had a higher word-stem completion rate than non-studied items in both the visual and auditory study modalities, which suggested activations of critical lures are not perceptual but perhaps conceptual in nature. In addition to failures to find modality-specific priming effects for critical lures (but see McKone & Murphy, 2000, for a report of modality specificity), there are studies where critical lure priming was not found (or was restricted) in the lexical decision task (McKone, 2004; Meade, Watson, Balota, & Roediger, 2007) and in the stem completion task when participants did not attempt to retrieve words from the study phase (McBride, Coane, & Raulerson, 2006).

Importantly, different theoretical hypotheses about the mechanisms of false memory activation underlie these contrasting views. The Implicit Associative Response hypothesis (IAR; e.g., Underwood, 1965; McDermott, 1997) proposes that the presence of a word (e.g., table) may activate the encoding of a related/critical word (e.g., chair). The IAR does not predict perceptual priming of critical lures directly. McKone and Murphy (2000) extended it to the IAR/imagery hypothesis (see also Hicks & Starns, 2005) in that the critical item might be encoded in the specific modality during the study phase (e.g., visual lists create visual (perceptual) representations of critical lures). Thus, critical lures are related to perceptual characteristics from a specific modality, which can explain why critical lures were effective in perceptual priming tasks. Alternatively, theories of false memory activation predict conceptual critical lure priming effects. The activation-monitoring (Roediger et al., 2001) and the associative-activation (Howe et al., 2009) theories suggest that processing of one concept activates a corresponding conceptual node and this activation spreads to surrounding associative concept nodes. When the DRM list items are presented, their shared associative concept – the critical lure – is most likely to be activated during encoding. Thus, these activation theories predict conceptual priming for critical lures as critical lures in the DRM paradigm primarily result from conceptual activation. However, it is also possible that perceptual features are stored along with the "nodes" in the associative network or they are integrated in the network in some way (Howe et al., 2009). For instance, false memories can be induced by phonologically related lists (e.g., *fat*, *cab*, *sat*, for the critical lure *cat*; Sommers & Lewis, 1999) and recently Finley, Sungkhasettee, Roediger, and Balota (2017) found that hybrid lists of conceptual and phonological items produced higher false memory rates than purely conceptual lists.

The main evidence that either supports or refutes the perceptual priming hypothesis for critical lures comes from data in priming tasks including the stem completion task (Hicks & Starns, 2005; McDermott, 1997; McKone & Murphy, 2000), the fragment completion task (McDermott, 1997), the anagram task (Lövdén & Johansson, 2003), the lexical decision task (McKone, 2004; Tse & Neely, 2005, 2007) and perceptual identification of degraded words (Hicks & Starns, 2005; Otgaar et al., 2015), all of which concern completing/identifying a word, (i.e., verbal in nature). We argue that there are two possible confounding factors in this previous research. First, the nature of these verbal priming tasks (perceptual vs. conceptual) is controversial in that some consider the stem completion task as conceptual not perceptual (Brown & Mitchell, 1994; Horton, Wilson, & Evans, 2001). Indeed, research has found that articulatory (verbal) suppression mediated critical lure priming in the stem completion task and in the anagram task (Lövdén & Johansson, 2003; Van Damme, Menten, & d'Ydewalle, 2010). For instance, when participants repeatedly said out loud the word "Coca-Cola" during encoding of DRM lists, critical lure priming disappeared. This suggests that verbal "perceptual" tasks might have a conceptual component, as purely perceptual tasks should not be affected by verbal suppression (Van Damme, Menten, & d'Ydewalle, 2010). To examine the nature of critical lures' memory traces, an unambiguous perceptual task needs to be used. Second, for verbal stimuli (i.e., words), it is difficult to completely dissociate perceptual information from conceptual meanings as words themselves usually constitute meanings (e.g., Can you communicate the meaning of "hot" without the word *hot*?). Thus, verbal stimuli might not be the best way to study the perceptual/conceptual nature of false memories' mental representations.

The Present Study

To deal with the above concerns, the present study used picture stimuli and a pictorial perceptual closure task to explore the consequences of false memories for visual stimuli and to test the perceptual/conceptual nature of false memories' mental representations. In Experiment 1, we evoked false memories for pictures by presenting pictorial versions of the DRM lists (see Israel & Schacter, 1997). Then, participants completed a recognition test to measure their false memory levels. Finally, participants received the pictorial PCT in which they were asked to identify degraded pictures of the studied (true) items as well as the critical lures as quickly as possible. Experiment 1 examined what consequences visual false memories might have on the identification performance of the PCT. In line with recent findings of the positive consequences of verbal false memories (Howe et al., 2010, 2011, 2013, 2015, 2016; Otgaar et al, 2015), one might expect to find that visual false memories are more easily recognized than visual true memories on the PCT. However, predictions can be made based on contrasting hypotheses about the nature of mental representation of false memories for visual stimuli. According to the dual coding theory, true memories for pictures are stored as both imaginal (perceptual) and verbal (conceptual) representations (Paivio, 1991). For instance, after seeing an image of a car, people will encode perceptual details such as shape and color of the car and they will also encode the concept “car”, which is verbal or linguistic. If false memories for visual stimuli also contain imagery representations, critical lure images should prime the pictorial PCT as fast as studied pictures. Conversely, if false memories for pictures are not perceptually based, false memories for visual stimuli might not prime the PCT as fast as presented pictures. In Experiments 2 and 3, we tested whether false memories for visual stimuli are conceptual (i.e., in a verbal form) in nature. We manipulated the modality of the study phase (picture vs. word) and modality of the PCT (picture vs. word). If false memory activations for critical pictures are conceptual, false memories for pictures should have a priming effect on the verbal PCT (i.e. identifying words).

Experiment 1

Before conducting Experiment 1, we conducted a pilot experiment with an identical procedure as in the current experiment. Based on the results from the pilot experiment, we have achieved effect sizes to estimate the sample size of Experiment 1 and we have obtained materials that are suitable for testing our hypotheses (for details see the supplementary materials).

Method

Participants

Based on the data from the pilot experiment (see supplementary materials), the Cohen’s d between the studied items and critical lures was from 0.52~ 0.76. To determine the sample size, we estimated the effect size as medium (d = 0.5). Using the software G*Power 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007), power analysis showed that thirty-four participants were needed when d was estimated as 0.5 and power was estimated as 0.80. Forty-five participants took part in this experiment. One participant was excluded due to misunderstanding of the instructions. Out of the 44 participants left, 6 were males and 38 females (M_{age} = 22.39, *SD* = 5.59) with age ranging from 19 to 54 years old. They received credit points or a financial compensation (€7.50). The experiment was approved by the ethical committee of the Faculty of Psychology and Neuroscience, Maastricht University.

Materials

Study phase. In the current experiment, nine DRM wordlists that could be represented pictorially were chosen from Peters, Jelicic, and Merckelbach (2015). Each DRM list included eight associated words (e.g., *truck*, *bus*, *train*, *jeep*, etc.) and these words were all related to a non-presented target or "critical lure" (i.e., *car*). In this experiment, we used pictorial versions of the DRM words (see also Israel & Schacter, 1997; Schacter, Israel, & Racine, 1999). The pictures were collected by means of a Google Images search. For example, for the word "*truck*", we searched for a picture of a "*truck*" using the Google Image search engine. All pictures were showing objects in the center (e.g., *car, bread*, *pen*) with a white background (see Appendix 3A for the DRM lists). To have a rough index that our chosen studied items and critical lures did not differ on certain critical dimensions, we looked at the following issues. First, we obtained the concreteness ratings (1 to 7) for the DRM list words (*M*= 6.0, *SD* = 0.55) and critical lures (*M*= 5.7, *SD* = 0.88) from Nelson, McEvoy, and Schreiber (1998) and found no significant difference between list words and critical lures ($p > .05$). Second, for the pictures of the DRM lists and critical lures, we asked a pilot sample of participants (n = 11, M_{age} = 24.1, *SD* = 4.66; 55% female) to name the objects in the pictures. 96.34% of the DRM pictures and 99.5% of the critical lure pictures were named correctly. Collectively, these analyses showed that our visual stimuli were concrete enough to be identified and would not lead to multiple interpretations. The pictures were presented digitally for 1000ms, with 1000ms inter-stimulus interval using E-Prime 2.0 software. Pictures within each list were presented in descending associative order.

Recognition test. The recognition test included 72 pictures in total. Of these pictures, 36 pictures were presented in the study phase (e.g., pictures of *truck, bus, train*, etc.; four randomly selected pictures per list), 9 were pictures of critical lures (e.g., image of *car*), 9 were not presented but related pictures (e.g., image depicting *bicycle*) and 18 were not presented and unrelated pictures (e.g. pictures of *tree*, *washing machine*, *glasses*, etc.). Each picture was shown briefly for 100 ms and there was a fixation cross on screen lasting 1500 ms after each picture.

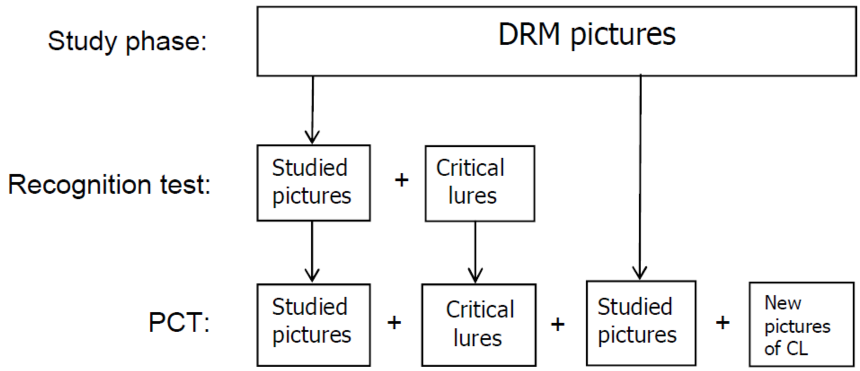


Figure 3.1. Diagram of procedure (CL: Critical Lures).

Perceptual closure task. The perceptual closure task (PCT) contained 108 pictures. Of all these pictures, 72 were studied pictures (36 were shown in the recognition test and 36 were not), 9 were pictures of critical lures from the recognition test, 9 were new pictures of the critical lures and 18 were unrelated pictures (9 were from the recognition test and 9 were not). As Figure 3.1 shows, for each kind of memory type (studied items, critical lures, unrelated), half of the pictures were shown during the recognition test and the other half was not. For instance, there were two sets of pictures for the critical lures. One set (9 pictures) were shown in the recognition test, the other set (9 new pictures of the critical lures) were not included in the recognition test but were tested directly in PCT (see Figure 3.1). For the 72 studied pictures, half were presented both in the DRM lists and in the recognition test while the other half were only shown in the DRM lists but not included in the recognition test. The pictures were displayed with a distortion filter that enabled the pictures to become clearer over time. We included 9 gradual gradations of blur, and the last one was the original picture (see Appendix 3B). Each blur was presented for 1 second and participants were asked to respond within 10 seconds. This task provides measures of both the accuracy and the reaction times for the identification of the pictures using the software E-Prime 2.0. The distorted pictures were created using GIMP 2.0.

Design and Procedure

A 3 (Memory type: studied items, critical lures, unrelated) × 2 (Recognition presentation: yes vs. no) within-subject design was used. Because the PCT is a priming task, the mere presentation of pictures in the recognition test might confound the results (but see Otgaar et al., 2015). In order to control the possible priming effect from the recognition test, a random half of the studied pictures in the PCT were presented in the recognition test and the other half were not shown during the recognition test. For the critical lures in the PCT, two versions of pictures for the same critical lure were prepared; one version was presented in the recognition test and the other was not. It was the same manipulation with unrelated pictures that half of the unrelated pictures were included in the recognition test and the other half were only shown during the PCT.

Participants were tested individually for approximately 45 minutes in lab facilities. First, participants were shown 9 pictorial DRM lists, with order of the lists randomized. Participants were instructed to memorize each picture and think about the concept of the picture when they saw it on the screen. After the presentation of all pictures, participants played Tetris for 5 min (filler task). Next, participants received the recognition test, during which pictures were subsequently presented and participants were asked to answer “yes” or “no” in a pop-up window to indicate whether they had seen the picture before or not. Finally, following five minutes of the filler task, participants received the perceptual closure task. They were asked to press the space bar as soon as they recognized the blurred picture. After pressing the button, a window popped up and asked participants to type in what they thought represented the degraded picture. No feedback was given regarding whether they had answered correctly or not. The pictures were successively presented in a pseudo- random order. Input results and reaction times were measured by the E-Prime 2.0 software.

Results and Discussion

The mean recognition rate for all studied pictures was 90.97% (*SD* = 9.15). The mean false recognition rate for critical lures was 26.83% (*SD* = 14.48), which is consistent with previous research using visual stimuli to induce false memories (see Israel & Schacter, 1997; Koutstaal, Schacter, & Brenner, 2001). Mean false recognition rate for unrelated items was 1.48% (*SD* = 3.80).

Before analyzing the data from the perceptual closure task, we filtered the data based on the following criterion (see also Otgaar et al., 2015). First, for studied pictures, we only included the reaction times for pictures that participants had recognized in the recognition test. This is to make sure they had formed true memories for these pictures. Second, for critical lures, only when participants falsely recognized the critical lures as presented during the recognition test were the reaction times for these pictures included in the analysis. This again, is to make sure they had formed false memories for the critical lure pictures. The above two criteria did not apply to pictures in the “no recognition presentation” condition as those pictures were not included in the recognition test. Third, pictures that were not identified correctly in the PCT were removed. This applied to all studied pictures, critical lures and unrelated items. For instance, if a participant pressed the button indicating identification of a picture (e.g., *butter*) but then did not write an answer or filled in an incorrect response (e.g., *chocolate*), then this was removed from the analyses. However, synonyms of the answers were accepted. For example, if a participant answered *marmalade* for a picture depicting *jam*, the answer was also counted as correct. Less than 5% of the accepted answers were synonyms. Of a total of 3484 answers that could be provided, 91.8% (N = 3200) were correct and 8.2% were blank or incorrect (N = 284).

We conducted a 3 (Memory type: studied items, critical lures, unrelated) ×2 (Recognition presentation: yes vs. no) repeated measures ANOVA to examine the PCT reaction time differences of the different memory types. Missing data (9 out of 264 cases) were replaced by the average value in the corresponding condition. As Figure 3.2 shows, no statistical interaction effect was found, $F(2, 86) = 1.63$, $p = .201$, $\eta^2_{partial} = .04$. There was a statistically significant main effect of memory type, $F(2, 86) = 27.51$, $p < .001$, $\eta^2_{partial} = .39$. Post-hoc Bonferroni tests revealed that studied items (M = 3011.63, SE =73.83) were identified statistically faster than critical lures (M = 3857.33, SE =136.89; $p < .001$) and unrelated items (M = 3678.92, SE =116.82; $p < .001$), and the latter two did not differ from each other ($p = .718$). We also found a significant main effect of recognition presentation, $F(1, 44) = 12.09$, $p = .001$, $\eta^2_{partial} = .22$.

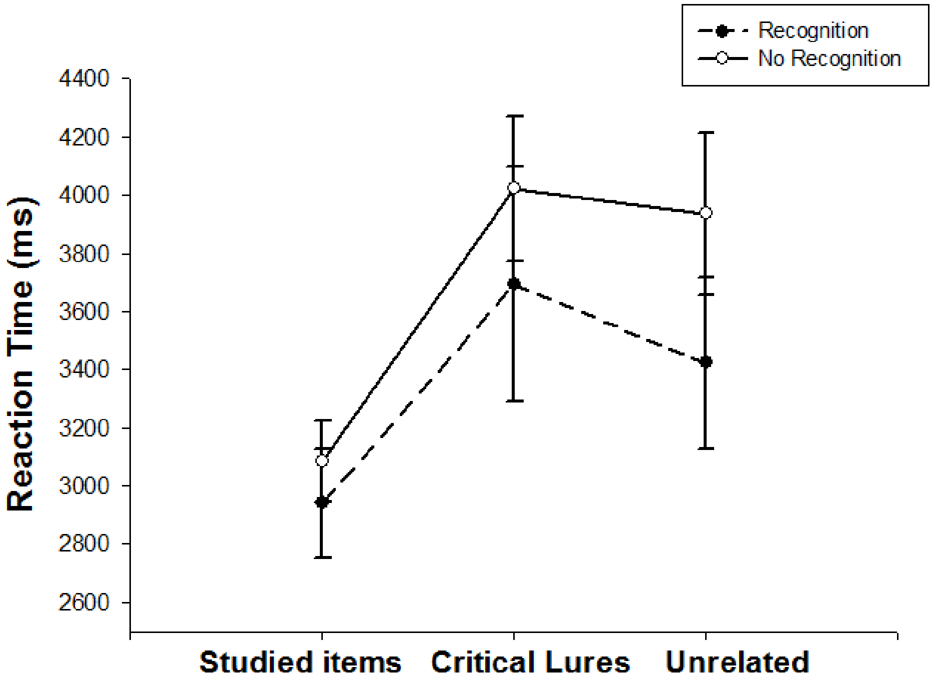


Figure 3.2. Mean reaction times in different Memory type and Recognition presentation conditions. Error bars indicate 95% confidence intervals (Experiment 1).

In the present experiment, we found that studied pictures (true memories) served as better primes for the PCT than false memories and the latter did not differ from unrelated pictures. That is, when participants were presented on the PCT with a picture of a non-presented critical item, they were statistically slower in identifying it than when they were presented with studied pictures. When using verbal stimuli (i.e., words), we and others (Howe et al., 2010; Otgaar et al., 2015) have, however, found that verbal false memories can lead to faster performance on tasks such as the PCT. Specifically, in these studies, it was found that false memories and not true memories served as the best primes. The current experiment shows that this finding cannot be translated to visual false memories or false memories that share more the characteristics of everyday experiences (e.g., color, details, etc.). False memories for pictures might not retain perceptual information like true memories do.

Our results on false recognition rates are in line with research on the mechanisms behind false recognitions for visual stimuli. In the current experiments, each picture was a visual analogue of a list member from the DRM list. Although pictures within one DRM list shared semantic relatedness with each other, each picture had distinctive perceptual characteristics. For instance, an image of a car differed in shape, color, and other details with a truck. Distinctive perceptual information is known to attenuate false memory and results in fewer pictures being false recognized (Arndt & Reder, 2003; Howe, 2008). We also found that stimuli presentation during recognition test improved reaction speed in the later identification task than when stimuli were not presented during the recognition test. However, stimuli exposure in the recognition test did not confound the priming effect of memory type.

Overall, we found that false memories for visual stimuli were less efficient primes on the PCT than visual true memories. We achieved the same results in our pilot experiment (see supplementary materials). The results led us to question the nature of mental representations of false memories induced by pictorial DRM lists. If false memories for pictures contain imagery representations as true memories do, retention of false memories should prime the identification of those images. Results from the current experiment did not support this hypothesis.

One might argue that the slower identification times for critical lure pictures were caused because studied pictures were presented more often (i.e., one time extra during the study phase) than critical pictures which were only presented during the test phase. To examine this hypothesis, we compared studied items that were not included in the recognition test (exposed once) with critical lures that were in the recognition test (also exposed once), and we found that critical lures were still more slowly identified than studied items, $t(43) = 3.43$, $p < .001$(see Figure 3.2). We also looked at the studied items that were only shown in the study phase (exposed once) and studied items that went through both study phase and recognition test (exposed twice). There was no significant difference in PCT performance, $t(43) = 1.96$, $p > .05$, indicating no effect of test on PCT performance. These analyses ruled out the possibility that our effects were due to the frequency of stimuli exposure. The most plausible explanation would be that mental representations for false memories derived from pictorial DRM lists are not visual, but verbal in nature. We suggest that false memories elicited by pictorial DRM lists are purely conceptual—not pictures but words in the mind’s eye.

According to spreading-activation theories (Howe et al., 2009; Roediger et al., 2001), presentation of DRM words activates corresponding concept nodes and this activation automatically spreads to interrelated nodes. Thus, non-presented but related concepts could be erroneously activated at the encoding phase. When it comes to presenting pictures of DRM list members, visual sensory traces are encoded for studied pictures and at the same time corresponding concepts of the pictures are activated automatically (based on the dual coding theory); however, such activation of the list members might only spread along the conceptual network as the DRM list members are semantically related instead of perceptually related. As a result, concepts of critical lures are activated and pictures of critical lures are falsely recognized during a recognition test. However, the activation of critical lure concepts might not work on the priming of critical lures for pictures. Previous studies have demonstrated the priming effect of false memories for verbal stimuli (McDermott, 1997; Otgaar et al., 2015; Tajika et al., 2005). The discrepancy between previous studies and our study led us to conduct a second experiment where we orthogonally manipulated the modality (picture vs. word) of both the DRM stimuli and the PCT. We aimed to test the hypothesis that mental representations of critical lures based on pictorial stimuli are conceptual, not perceptual or imaginal, in nature.

Experiment 2

The purpose of Experiment 2 was to examine the priming effect of false memories for both visual and verbal stimuli on the PCT in one study. Previous study showed that verbal false memories primed the verbal PCT faster than true memories (Otgaar et al., 2015). Our prediction was that we would find this effect in Experiment 2 and we predicted that the reverse was true for false memories for visual stimuli, i.e. false memories for visual stimuli primed the PCT slower than true memories. More importantly, we were interested in whether false memories for visual stimuli have priming effects on identification of words. Experiment 2 was an orthogonal design with manipulation of verbal vs. visual modalities on both study phase and the PCT.

Method

Participants

The Cohen's *d* between studied items and critical lures was 0.68 in the recognition presentation condition and 1.42 in the no recognition presentation condition in Exp. 1. Again, we were interested in the reaction time contrast between studied items and critical lures in each group. We estimated *d* = 0.7 and power as 0.8, G. Power analysis showed nineteen participants were needed in each group and thus 76 participants were needed in total. Eighty participants were tested in Experiment 2. The mean age of the sample was 20.6 years old (*SD* = 2.20; age range = 18 to 29 years old). There were 8 males and 72 females. They received credit points or a financial compensation (€ 7.50).

Materials

Study phase. The nine DRM lists (eight-item per list) from Experiment 1 were used. However, Experiment 2 used two sets of stimuli: verbal DRM words and pictorial version of the verbal DRM lists. The pictorial version was exactly the same as used in Experiment 1.

Recognition test. Recognition items were the same as in Experiment 1, except in Experiment 2 a verbal version of the recognition items was used as well. In the verbal version, each word (e.g., *sandwich*) was a corresponding concept to a picture in the pictorial version (e.g., a picture of a sandwich).

Perceptual closure task. The PCT had two versions, a word version and a picture version. The two versions contained the same conceptual items of studied and critical items, only with different presentation forms. For instance, the item "car" in the word version of the PCT was the word *car*, while in the picture version was a picture of a car. The filtering parameters (level of degrading) for words were the same as the filters for pictures in Experiment 1. The word version of the PCT included 99 words, consisting of 72 studied words (36 were shown in the recognition test and 36 were not), 9 critical lures, 9 related words and 9 unrelated words. The picture version of the PCT was composed of 99 pictures, including 72 studied pictures (36 were shown in the recognition test and 36 were not), 9 pictures of critical lures from the recognition test, 9 new pictures of the critical lures and 9 unrelated pictures.

Design and Procedure

Experiment 2 involved a 2(DRM modality: picture vs. word) x 2(PCT modality: picture vs. word) x 3(Memory: studied items, critical lures, unrelated items) mixed design where the first and second factors were between-subjects and the third factor was within-subject. Participants were tested individually for approximately 40 minutes in lab facilities. The procedure was exactly the same as in Experiment 1, but we manipulated the stimuli modality (picture vs. word) in the study phase and in the PCT phase. Thus, we had four groups: (1) picture-picture group, where DRM lists were presented in a pictorial form in the study phase and blurred pictures had to be identified in the PCT; (2) picture-word group, where pictorial DRM items were presented but blurred words for these items had to be identified in the PCT; (3) word-word group, words were presented and later identified in the PCT; and (4) word-picture group, where DRM words were presented in the study phase and then pictorial forms had to be identified in the PCT. Participants were randomly assigned to one of the four groups. To be noted, the modality of the recognition test was always consistent with the study phase.

Results and Discussion

The mean recognition rates in different groups of participants are shown in Table 3.1. Eighty participants were equally distributed in the four groups. Inspection of Table 3.1 shows that when the DRM lists were presented verbally, the false recognition rates for critical lures were higher than when they were presented pictorially ($F(1, 78) = 40.88$, $p < .001$). However, this discrepancy is not a confounding factor to the PCT results. This is because the data analysis is not contingent on differences in true or false memory rates, just that a true or false memory had been generated during the recognition test (i.e., we only analyse the PCT data for which true or false recognition occurred). Before analyzing the data from the PCT, we conducted the same filtering process as in Experiment 1. That is, first, we focused only on the reaction times of recognized studied and critical items and then, within these recognized items, items with incorrect answers were removed from the analysis. Of a total of 2595 recognized items, for 10.8% (N=280) items participants failed to identify them within 10 s, thus no answers and reaction times were recorded by the program; for 5.4% (N=141) of the items, participants provided incorrect answers, of which the reaction times were removed from analysis. Incorrect answers in unrelated items were also removed. Of a total of 720 unrelated items, participants failed to identify 14.6% (N=105) of the items, and answered incorrectly for 11.1% of the items (N=80). Those items were not included in analysis.

Table 3.1

Mean recognition rates for studied items, critical lures and unrelated items in four groups of participants (n: number of participants in that condition)

	Picture-Picture (n=20)	Picture-Word (n=20)	Word-Picture (n=20)	Word-Word (n=20)
Studied items	87.86%	87.14%	77.43%	79.14%
Critical lures	26.11%	23.89%	52.22%	50.00%
Unrelated items	1.11%	0.56%	6.11%	3.89%

The main interest was to compare PCT reaction times to studied items and critical lures in each group, so we conducted a 4(Group: picture-picture, picture-word, word-picture, word-word) x 3(Memory: studied items, critical lures, unrelated items) repeated measures ANOVA, with Group as a between-subjects variable. Because Experiment 1showed that stimuli exposure in the recognition test did not interact with the reaction time contrast among studied items, critical lures and unrelated items, the items for which reaction times were analyzed below were all presented in the recognition test. Missing data (4 out of 240 cells) were replaced by the average value in the corresponding condition. A statistically significant interaction effect between Group and Memory was found, $F(6, 152) = 8.01$, $p < .001$, $\eta^2_{partial} = .24$ (see Figure 3.3). There was a main effect of Memory, $F(2, 152) = 19.25$, $p < .001$, $\eta^2_{partial} = .20$, and a main effect of Group, $F(3, 76) = 3.86$, $p = .013$, $\eta^2_{partial} = .13$. Simple effect analyses indicated the following: (1) in the picture-picture group, studied items (M= 3268, SD=676) primed the PCT significantly faster than critical lures (M= 4069, SD=1815; $p = .047$, Cohen's $d = 0.57$) and unrelated items (M= 3558, SD=780; $p = .03$, Cohen's $d = 0.55$), and the latter two did not differ from each other (p= .25), which is consistent with findings in Experiment 1; however, (2) critical lures (M= 3595, SD= 848) primed the PCT significantly faster than studied items (M= 4309, SD=635) in the picture-word group ($p < .001$, Cohen's $d = 0.99$), and both were statistically faster than the unrelated items (M= 5027, SD= 517; ps < .001, Cohen's $d_{\text{studied items}}$= 1.40, Cohen's $d_{\text{critical lures}}$= 2.28); (3) in the word-picture group, critical lures (M= 3731, SD=1073) and studied items (M= 3873, SD= 567) did not differ from each other ($p = .58$), and both did not differ from unrelated items (M= 4217, SD=759; ps > .05) in priming the PCT; and (4) in the word-word group, critical lures (M= 3228, SD= 904) were significant faster primes than studied items in the PCT (*M*= 3720, *SD*= 716; p = .02, Cohen's d = 0.62), and both were faster than unrelated items (*M*= 4622, *SD*= 747; *p*s < .001, Cohen's $d_{\text{studied items}}$= 1.31, Cohen's $d_{\text{critical lures}}$= 1.68), which replicated the findings by Otgaar et al. (2015).

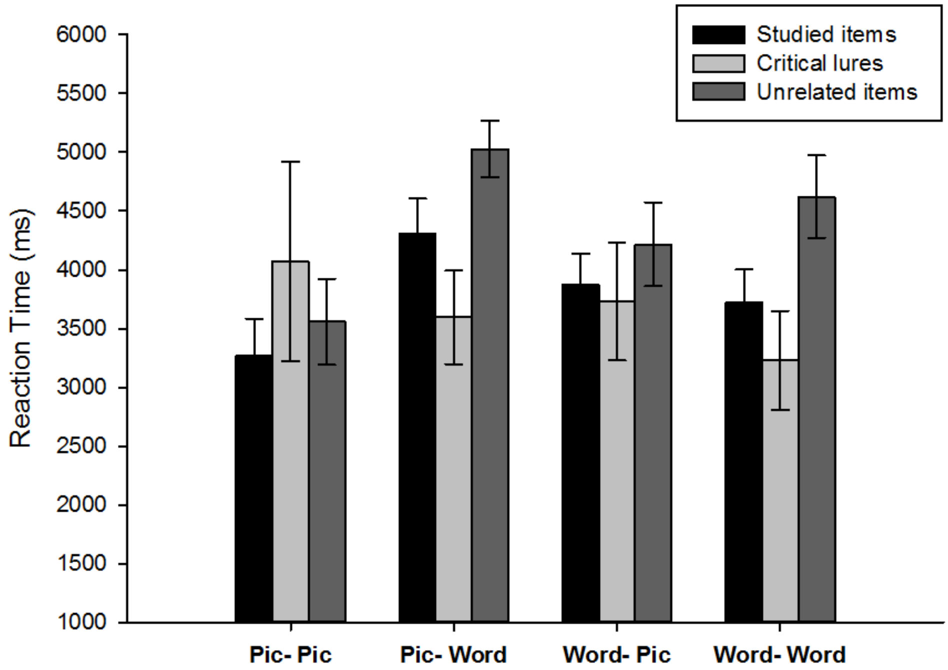


Figure 3.3. Mean reaction times to studied items, critical lures and unrelated items in different study/PCT modalities ("Pic" stands for "Picture"). Error bars indicate 95% confidence intervals (Experiment 2).

Experiment 2 replicates the results of Experiment 1 that false memories induced by pictorial DRM lists primed the pictorial PCT slower than true memories did, and they were no significantly different from unrelated pictures. Participants falsely recognized critical lure pictures as studied in the recognition test, however false memory activation did not have any facilitation on identifying critical lure pictures. Memory activation of studied pictures, on the other hand, improved the speed in the identification of studied pictures. Experiment 2 also replicates the finding by Otgaar et al. (2015) that verbal false memories induced by verbal DRM lists had a priming effect on the PCT, and this effect was even more superior than that of verbal true memories.

Experiment 2 also provided two new findings. First, we found that both false and true memories induced by pictorial DRM lists had a priming effect in the verbal PCT, and the magnitude of false memories' priming effect was larger than that of true memories for pictorial DRM lists. This result pattern is the same as in the word-word condition in this experiment, where verbal DRM lists were presented and verbal PCT was completed. It can be best explained by the assumption that false memories elicited by DRM pictures are verbal in nature. When pictures of DRM items were presented, concepts of these items were activated and spread to nearby concepts, i.e. critical lures, in a semantic network. The activation of critical lure concepts later facilitated the identification of critical lure words. If DRM pictures created *visual* false memories (i.e. vivid imagery representations of critical lure pictures), then visual false memories should have a priming effect in the pictorial PCT, which was not found in our two experiments. Thus, false memories from pictorial DRM lists should be encoded in other forms, and the Picture-Word dataset in Experiment 2 supports the verbal form.

Second, we found that in the word-picture condition, where DRM words were presented and pictorial forms were identified in the PCT, neither false memories nor true memories had a priming effect. This is logical as in the study phase both verbal true memories and verbal false memories were created. According to the spreading-activation theories, conceptual nodes for studied items and critical lures are embedded in one semantic network and are both activated during study, which means the encodings of true and false memories are both verbal. As a result, verbal true memories and verbal false memories should not prime identification of pictures in the PCT.

A limitation of Experiment 2 is that we asked participants to think about the concepts of the pictures during the presentation of pictorial DRM lists. This instruction was borrowed from Israel and Schacter (1997) in order to maximize the chances of obtaining false memories of critical lures within the DRM paradigm. Nevertheless, it might have switched participants' attention from viewing the pictures to conceptually processing the pictures. Experiment 3 was conducted to rule out this possible confounding factor. Here, participants were merely asked to carefully view the presented pictures/words during the study phase of Experiment 3. Moreover, data on Image Agreement Ratings (Snodgrass & Vanderwart, 1980) of stimuli were obtained and reported in Experiment 3.

Experiment 3

Method

Participants

An a priori sample size estimation was conducted as per Experiment 2, and suggested that at least 76 participants were necessary. We tested 84 participants in Experiment 3 (M_{age} = 20.18, *SD* = 2.76; age range = 18 to 40 years old; 83% female). Participants received credit points for participation in the experiment.

Materials

The same verbal and pictorial DRM lists (eight-item per list) from Experiment 2 were used. For the orientation of the DRM pictures, we used: (1) objects standing in their natural (vertical) positions if possible (the *car* list, the *bread* list, the *fruit* list, the *shirt* list, the *cup* list, the *chair* list and the *foot* list); (2) long, thin objects that cannot stand (the *needle* list and the *pen* list) oriented at 45 degree angle (upper right to lower left). Critical lure pictures were always in the same orientation with most members of the corresponding studied list. We administered an Image Agreement Ratings questionnaire to 43 participants (M_{age} = 20.19, *SD* = 1.97; 88% female), asking them to rate to what extent the DRM picture represents their mental image of the DRM word on a 1 to 5 scale (1 = no agreement; 5 = high agreement). Of 54 pictures whose PCT data would be entered into data analysis, three pictures (two studied pictures and one unrelated pictures) had an average imagery agreement rating lower than 3. Thus, PCT data of these three pictures were not entered into analyses. After removing these three pictures, a repeated measures ANOVA was conducted to compare the mean imagery agreement ratings among studied pictures (*M* = 4.22, *SD* = 0.48), critical pictures (*M* = 4.23, *SD* = 0.54) and unrelated pictures (*M* = 4.25, *SD*= 0.46). No statistical difference on imagery agreement ratings was found among different type of stimuli, $F(2, 84) = 0.16$, $p = .85$, $\eta^2_{partial} = .004$.

Design and Procedure

Experiment 3 used the same 4 (Group: picture-picture, picture-word, word-picture, and word-word) x 3 (Memory: studied items, critical lures, unrelated items) mixed design as in Experiment 2, where Group was between-subjects and Memory was within-subject. The procedure was identical to that of Experiment 2, except the instruction in the study phase was modified such that in Experiment 3 participants were merely asked to view the pictures carefully.

Results and Discussion

Mean recognition rates of studied items and critical lures in different groups are shown in Table 3.2. Participants in Experiment 3 had similar false memory rates for critical pictures (~35%) with previous experiments (25%~30%). This suggests that telling participants to think about the concept of the presented picture or not did not impact false recognition of critical pictures.

Table 3.2

Mean recognition rates for studied items, critical lures and unrelated items in four groups of participants (n: number of participants in that condition)

	Picture-Picture (n=20)	Picture-Word (n=22)	Word-Picture (n=22)	Word-Word (n=20)
Studied items	87.95%	87.00%	79.22%	74.45%
Critical lures	35.63%	35.86%	59.66%	40.00%
Unrelated items	1.11%	3.54%	4.55%	3.33%

After filtering wrong answers and non-recognized studied items and critical pictures as in previous experiments, a 4 (Group: picture-picture, picture-word, word-picture, word-word) x 3 (Memory: studied items, critical lures, unrelated items) repeated measures ANOVA was conducted, with Group as a between-subjects variable. A statistically significant interaction effect between Group and Memory was found, $F(6, 152) = 4.26$, $p = .001$, $\eta^2_{partial} = .14$. There was a main effect of Memory, $F(2, 152) = 23.06$, $p < .001$, $\eta^2_{partial} = .23$, and a main effect of Group, $F(3, 76) = 10.84$, $p < .001$, $\eta^2_{partial} = .30$.

As can be seen in Figure 3.4, and verified using simple effect analyses, the following effects were obtained: (1) in the picture-picture group, studied items (M = 3005, SD = 444) primed the PCT statistically faster than critical lures (M = 3632, SD = 1249; p = .036, Cohen's d = 0.65) and unrelated items (M = 3799, SD = 830; p = .001, Cohen's d = 0.98), and the latter two did not differ from each other (p = .64), which is consistent with findings in Experiments 1and 2; however, (2) in the picture-word group, critical lures (M = 3879, SD = 911) primed the PCT significantly faster than studied items (M = 4371, SD = 548) (p =

.005, Cohen's d = 0.75), and both were statistically faster than the unrelated items (M = 5188, SD = 462; ps < .001, Cohen's $d_{\text{studied items}}$ = 1.54, Cohen's $d_{\text{critical lures}}$ = 1.62), which again replicated results of Experiment 2; (3) in the word-picture group, studied items (M = 3728, SD = 450) and critical lures (M = 3673, SD = 986) did not differ from each other (p = .81), and neither differed from unrelated items (M = 4081, SD = 890; ps > .05) in priming the PCT; and (4) in the word-word group, critical lures (M = 4188, SD = 1361) and studied items (M = 3672, SD = 611) did not statistically differ in priming the PCT (p = .10), and both were faster than unrelated items (M = 4678, SD = 611; ps < .05, Cohen's $d_{\text{studied items}}$ = 1.74, Cohen's $d_{\text{critical lures}}$ = 0.64).

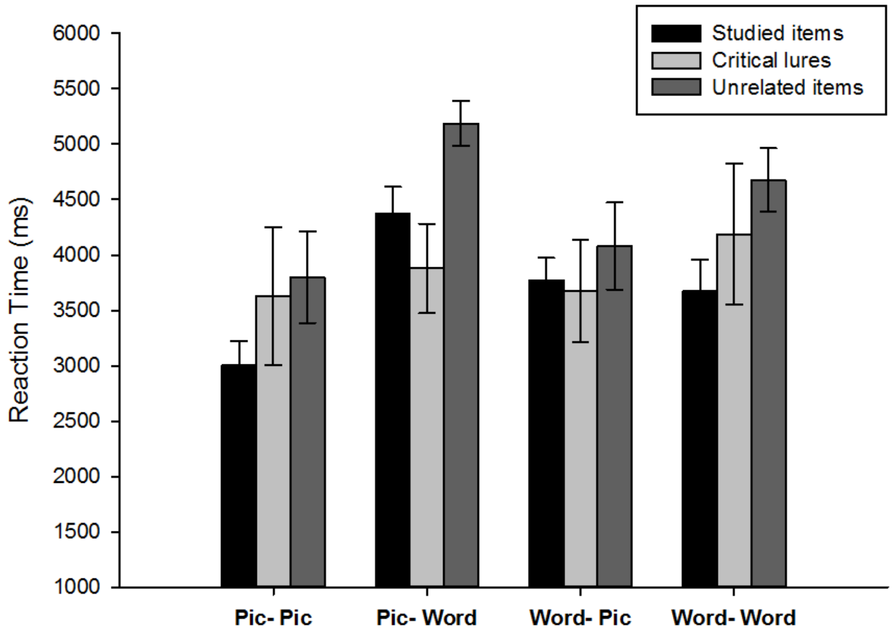


Figure 3.4. Mean reaction times to studied items, critical lures and unrelated items in different study/PCT modalities ("Pic" stands for "Picture"). Error bars indicate 95% confidence intervals (Experiment 3).

Experiment 3 replicated the results of Experiment 2. The most relevant results are those of the picture-picture and picture-word groups. In these two conditions of Experiment 3, participants were instructed to view the pictures carefully, which should direct them to perceptually processing the pictures. However, without telling participants to think about the concepts of DRM pictures, false memories of critical pictures still primed identification of corresponding concepts more quickly than true memories of studied pictures (picture-word condition) but lacked a priming effect on identification of the exact same critical pictures (picture-picture condition). The results of Experiment 3 provide converging support for our hypothesis that false memories for visual stimuli are conceptual in nature.

General Discussion

Our experiments focused on the consequences of false memories on an adapted perceptual closure task. We investigated the extent to which false memories were better or worse primes on the PCT than true memories. After presenting participants with visual DRM stimuli, the majority formed false memories for non-presented pictures of critical lures. In all three experiments, we found that these false memories did not result in any priming effect in the PCT relative to unrelated pictures, while visual true memories exhibited facilitation in priming the PCT. More specifically, perceptual priming based on visual true memories was faster than that based on false memories for critical lure pictures. However, when we changed the modalities of the study phase and the PCT phase in Experiments 2 and 3, we found that false memories for critical lure pictures resulted in a larger priming effect than true memories in the verbal PCT.

Consistent with the spreading-activation theories, our results support the hypothesis that false memory representations do not contain perceptual information. Although previous studies found that verbal false memories can have priming effects on verbal priming tasks (e.g., the stem completion task, McDermott, 1997; McKone & Murphy, 2000), these tasks might include a component of conceptual priming and thus, it is not fully justified to interpret them as evidence for perceptual priming for false memories. In fact, Hicks and Starn’s (2005) finding of no modality specificity for priming in the stem completion task provides evidence for the conceptual priming in the stem completion task. However, with only verbal priming tasks, Hicks and Starn’s study (2005) cannot provide direct evidence regarding whether false memory representation is perceptual or not. By using pictorial stimuli, our study showed that identification of degraded pictures was not facilitated by critical lure pictures, which provides strong and direct evidence for the non-perceptual view of false memory representations.

We consistently found in three experiments that false memories for pictures primed the PCT more slowly than true memories for pictures. However, in the picture-word group, we found that false memories for pictures had a priming effect in the verbal PCT. To explain the contrasting findings, our interpretation is that memory representations underlying false memories for visual stimuli might be purely verbal or conceptual. Theoretically, this can be explained as follows. The dual coding theory posits that true memories for visual stimuli have two forms of memory coding, imaginal and verbal (Paivio, 1991). In the study phase, true memories for DRM pictures (e.g., a picture of an apple) are encoded as imaginal representations (e.g., image of an apple) as well as verbal representations (e.g., the concept “apple”). The activation of verbal representations is automatic according to the dual-coding theory, and this activation spreads through a semantic network which activates nearby concepts of critical lures (Howe et al., 2009; Roediger et al., 2001). Thus, true memories for visual stimuli are encoded into imaginal and verbal forms while false memories for visual stimuli are only encoded into verbal forms. This interpretation can indeed explain our findings. Since true memories for pictures possess imaginal representations while false memories do not, true memories primed identification of pictures faster than false memories in the pictorial PCT. Since false memories for pictures possess conceptually-based representations like true memories do, critical lure pictures had a priming effect on the verbal PCT.

Interestingly, the above-mentioned explanation is in line with Fuzzy-trace Theory (Brainerd & Reyna, 2002; Brainerd, Reyna, & Ceci, 2008). According to the Fuzzy-trace Theory, two independent memory traces, verbatim and gist, are formed when people encode stimuli. Verbatim traces refer to episodic representations of the surface forms of experienced items, which correspond to the perceptual representations. Gist traces are interpretations of concepts (e.g., meanings) that are retrieved from the surface forms, which correspond to the conceptual aspect of mental representations. Fuzzy-trace Theory proposes that both verbatim (or perceptual) traces and gist traces contribute to true memories but only gist traces contribute to the formation of false memories. Related to our results, true memories of pictures thus primed identification of both blurred pictures and words while false memories of pictures only primed identification of words.

Our hypothesis is consistent with neuroimaging findings on brain activity differences between true and false memories for visual stimuli. In Slotnick and Schacter’s (2004, 2006) studies, participants studied sets of exemplar shapes and formed false memories for non-presented but related shapes. Researchers found that true recognition for images revealed more activities in the early visual cortex compared with false recognition. Stark, Okado, and Loftus (2010) also found stronger activation for visual true memories in early visual processing areas, which are mostly concerned with sensory processing of visual details. Schacter, Chamberlain, Gaesser, and Gerlach (2012) suggest that true memories are accompanied by reactivation of sensory encoding processes that were engaged during the establishment of true but not false memories. We posit that true memory for visual stimuli contain imaginal and verbal representations while false memory for visual stimuli possesses only verbal representations, at least when they are evoked in the DRM paradigm. Activation in early visual areas probably corresponds to the imaginal representation for visual true memory, which can provide a neural basis for the differentiation of memory encoding between true and false memories for visual stimuli.

Another interesting finding is that in the picture-word and word-word conditions, false memories sometimes primed the PCT even faster than true memories. This finding can be explained by the principle of additivity in the spreading-activation theory of false memory. The principle of additivity suggests that activation of the critical lure can summate to produce greater priming effect under multiple sources of activation (e.g., multiple DRM studied items) than under a single source of activation (e.g., a studied item) (Roediger et al., 2001; Meade, Hutchison, & Rand, 2010). That is, during the study phase each concept of the DRM list was probably activated once while the concept of the critical lure had more chance of activating multiple times. Thus, false memories primed the identification of concepts faster than true memories.

The finding of false memories' superior priming effect demonstrates the adaptive nature of memory as well. In recent years, studies have accumulated showing that false memories have adaptive value and can exhibit positive consequences under certain conditions (Howe, 2011; Schacter, 2012). By "positive consequences" it means that false memories show the same efficiency as true memories or even exert more efficiency (i.e., faster identification) than true memories in solving PCT problems. It has been well established by previous studies that verbal false memories can have salutary consequences. Verbal false memories are found to be associated with convergent thinking--a component of creativity (Dewhurst, Thorley, Hammond, & Ormerod, 2011). Recent studies have confirmed that false memories derived from DRM paradigm solved the PCT faster, solved associative problems faster, and solved analogical problems faster than true memories

(Howe, Garner, & Patel, 2013; Otgaar et al., 2015; Wilkinson, Howe, & Monaghan, 2018). In our experiments, we replicated previous findings in the word-word group and added another new condition (i.e., picture-word) where false memories exhibit positive consequences as well. Our study suggests that, as long as true and false memories share the same memory activation/representation mechanism (e.g., derived from the same semantic memory network), false memories can have beneficial effects on related tasks (e.g., a semantic associative task).

Some caveats are worth mentioning. In previous studies, participants reported vividly seeing pictures when words were shown and sometimes even seeing pictures that were not shown (Foley, Bays, Foy, & Woodfield, 2015; Foley & Foy, 2008; Weinstein & Shanks, 2010). These false memories were reported with high confidence and a feeling of recollection. Indeed, false memories are caused by different types of associations (e.g., conceptual, perceptual; Howe et al., 2009; Pesta, Murphy, & Sanders, 2001). Although our data in the word-picture condition did not support the perceptual priming of words on identifying corresponding pictures, we were using the DRM paradigm in all three experiments, which might be a limitation of our study. For instance, Weinstein and Nash (2013) found perceptual priming for false memories using a visual scene scenario, although they did not control the recognition presentation and their effect size was not large. Moreover, we asked participants to “view the pictures carefully” to guide their attention to perceptually processing the stimuli. Future studies may consider using encoding instructions that focus even more on processing perceptual elements such as color and shape. Further research is needed to examine the underlying representation of false memories by using other false memory paradigms and tasks.

One might ask whether results in our experiments were due to item effects. Data from previous PCT research and data from the current experiments indicate that this is probably not the case. In the verbal PCT study of Otgaar et al. (2015) and associative problem-solving tasks by Howe et al. (e.g., 2010, 2016), the issue of possible item effects has been handled by comparing presented critical lures and non-presented critical lures. These studies still found a priming effect for non-presented critical lures, which is consistent with our results in the word-word condition. As for the results in the picture-picture condition, the pictures in Experiment 1 and our pilot experiment were different stimuli, but we achieved the same results in the two experiments, which eliminates the possibility that our results were due to the use of specific materials. Furthermore, DRM lists in three Experiments were all about familiar objects. We have controlled for the concreteness, name agreement, and image agreement of studied items and critical lures: a pilot name agreement study indicated that both studied pictures and critical pictures were easy to identify (both > 96%, see *Materials* of Exp. 1); and an image agreement ratings questionnaire showed that participants' mental images of DRM items were in high accordance with the pictures used in our experiments (~ 4.2 in a 1-5 scale, see *Materials* of Exp. 3).

Some might argue that the different performance between visual true and false memories on the PCT is due to mere stimuli exposure: studied pictures were presented more often than the critical lures and unrelated pictures (during study phase). What we want to argue here is that we only included reaction times for pictures in which participants responded with "yes" during the recognition test. For studied pictures, only pictures they had recognized in the recognition test were entered into analysis to make sure we were analysing true memories. The same is true for critical lures that only pictures that were falsely remembered were analysed. This means that the comparison level is not stimuli exposure versus no stimuli exposure, but true versus false memory. Furthermore, we examined whether stimuli exposure could explain our effects but we did not find any evidence for this idea. What we want to stress here is that, unlike verbal false memories, false memories for visual stimuli are processed in a way not different from unrelated pictures, and both are more slowly processed than visual true memories. Research on false memories relies heavily on paradigms using verbal stimuli. Studying false memories for stimuli with more ecological validity such as images and events might contribute to a more complete picture of false memory mechanism(s).

To summarize, our experiments provide a first examination of the perceptual priming effect of false memories for visual stimuli in the DRM paradigm. We found that false memories for pictures did not prime identification of degraded pictures, but they had a priming effect on identifying blurred words. Based on these results, we suggest that true and false memories differ in the way they are encoded in that true memories for visual stimuli have both verbal and imaginal memory traces while false memories for visual stimuli are only encoded in a verbal form. Our results contribute to a more complete picture about the differentiation between true and false memories and have implications concerning the underlying mechanisms of false memories.

References

Arndt, J., & Reder, L. (2003). The effect of distinctive visual information on false recognition. *Journal of Memory and Language, 48*, 1–15.

Brown, A. S., & Mitchell, D. B. (1994). A reevaluation of semantic versus nonsemantic processing in implicit memory. *Memory & Cognition*, *22*, 533-541.

Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science, 11*, 164-169.

Brainerd, C. J., Reyna, V. F., & Ceci, S. J. (2008). Developmental reversals in false memory: A review of data and theory. *Psychological bulletin, 134*, 343-382.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*, 17–22.

Dewhurst, S. A., Thorley, C., Hammond, E. R., & Ormerod, T. C. (2011). Convergent, but not divergent, thinking predicts susceptibility to associative memory illusions. *Personality and Individual Differences, 51*, 73–76.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175-191.

Finley, J. R., Sungkhasettee, V. W., Roediger, H. L., & Balota, D. A. (2017). Relative contributions of semantic and phonological associates to over-additive false recall in hybrid DRM lists. *Journal of Memory and Language, 93*, 154-168.

Foley, M. A., Bays, R. B., Foy, J., & Woodfield, M. (2015). Source misattributions and false recognition errors: Examining the role of perceptual resemblance and imagery generation processes. *Memory, 23*, 714–735.

Foley, M. A., & Foy, J. (2008). Pictorial encoding effects and memory confusions in the Deese-Roediger-McDermott Paradigm: Evidence for the activation of spontaneous imagery. *Memory, 16*, 712–727.

Hicks, J. L., & Starns, J. J. (2005). False memories lack perceptual detail: Evidence from implicit word-stem completion and perceptual identification tests. *Journal of Memory and Language*, *52*, 309-321.

Horton, K. D., Wilson, D. E., & Evans, M. (2001). Measuring automatic retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 958-966.

Howe, M. L. (2006). Developmentally invariant dissociations in children's true and false memories: Not all relatedness is created equal. *Child Development, 77,* 1112–1123.

Howe, M. L. (2008). Visual distinctiveness and the development of children's false memories. *Child Development, 79,* 65–79.

Howe, M. L. (2011). The adaptive nature of memory and its illusions. *Current Directions in Psychological Science, 20,* 312–315.

Howe, M. L., & Derbish, M. H. (2010). On the susceptibility of adaptive memory to false memory illusions. *Cognition, 115*, 252–267.

Howe, M. L., Garner, S. R., Dewhurst, S. A., & Ball, L. J. (2010). Can false memories prime problem solutions? *Cognition, 117,* 176–181.

Howe, M. L., Garner, S. R., Charlesworth, M., & Knott, L. M. (2011). A brighter side to memory illusions: False memories prime children's and adults' insight-based problem solving. *Journal of Experimental Child Psychology, 108*, 383–393.

Howe, M. L., Garner, S. R., & Patel, M. (2013). The positive consequences of false memories. *Behavioral Science and the Law, 31*, 652-665.

Howe, M. L., Garner, S. R., Threadgold, E., & Ball, L. J. (2015). Priming analogical reasoning with false memories. *Memory & Cognition*, *43*, 879-895.

Howe, M. L., & Otgaar, H. (2013). Proximate mechanisms and the development of adaptive Memory. *Current Directions in Psychological Science*, *22,* 6–22.

Howe, M. L., Threadgold, E., Norbury, J., Garner, S., & Ball, L. J. (2013). Priming children's and adults' analogical problem solutions with true and false memories. *Journal of Experimental Child Psychology*, *116,* 96-103.

Howe, M. L., Wilkinson, S., Garner, S. R., & Ball, L. J. (2016). On the adaptive function of children's and adults' false memories. *Memory*, *24*, 1062–1077.

Howe, M. L., Wimmer, M. C., Gagnon, N. & Plumpton, S. (2009). An associative activation theory of children's and adults' memory illusions. *Journal of Memory and Language*, *60,* 229-251.

Israel, L., & Schacter, D. L. (1997). Pictorial encoding reduces false recognition of semantic associates. *Psychonomic Bulletin and Review, 4,* 577–581.

Koutstaal, W., Schacter, D. L., & Brenner, C. (2001). Dual task demands and gist-based false recognition of pictures in younger and older adults. *Journal of Memory and Language, 44*, 399 – 426.

Lövdén, M., & Johansson, M. (2003). Are covert verbal responses mediating false implicit memory? *Psychonomic Bulletin & Review, 10*, 724-729.

Luteijn, F., & Barelds, D. P. H. (2004). *GIT2 Groninger Intelligentie Test 2. Handleiding*. Amsterdam: Harcourt Test Publishers.

McBride, D. M., Coane, J. H., & Raulerson, B. A. (2006). An investigation of false memory in perceptual implicit tasks. *Acta Psychologica, 123*, 240-260.

McDermott, K. B. (1997). Priming on perceptual implicit memory test can be achieved through presentation of associates. *Psychonomic Bulletin and Review, 4*, 582–586.

McKone, E. (2004). Distinguishing true from false memories via lexical decision as a perceptual implicit test. *Australian Journal of Psychology, 56*, 42-49.

McKone, E., & Murphy, B. (2000). Implicit false memory: effects of modality and multiple study presentations on long-lived semantic priming. *Journal of Memory and Language, 43*, 89–109.

Meade, M. L., Hutchison, K. A., & Rand, K. M. (2010). Effects of delay and number of related list items on implicit activation for DRM critical items in a speeded naming task. *Journal of Memory and Language, 62*, 302-310.

Meade, M. L., Watson, J. M., Balota, D. A., & Roediger, H. L. (2007). The roles of spreading activation and retrieval mode in producing false recognition in the DRM paradigm. *Journal of Memory and Language, 56*, 305-320.

Miller, B., & Gazzaniga, M.S. (1998). Creating false memories for visual scenes. *Neuropsychologia*, *36*, 513–520.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. http://w3.usf.edu/FreeAssociation/.

Otgaar, H., Sauerland, M., & Petrila, J. (2013). Novel shifts in memory research and their impact on the legal process. *Behavioral Sciences and the Law, 31*, 531–540.

Otgaar, H., Howe, M. L., Peters, M., Smeets, T., & Moritz, S. (2014). The production of spontaneous false memories across childhood. *Journal of Experimental Child Psychology, 121*, 28-41.

Otgaar, H., Howe, M. L., van Beers, J., van Hoof, R., Bronzwaer, N., & Smeets, T. (2015). The positive ramifications of false memories using a perceptual closure task. *Journal of Applied Research in Memory and Cognition, 4*, 43–50.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology, 45*, 255-287.

Perters, M. J. V., Jelicic, M., & Merckelbach, H. (2015). *Inducing false memories: A Dutch version of the Deese/Roediger-McDermott paradigm.* Unpublished manuscript.

Pesta, B. J., Murphy, M. D., & Sanders, R. E. (2001). Are emotionally Charged Lures immune to False Memory? *Journal of Experimental Psychology: Learning, Memory and Cognition*, *27*, 328-338.

Roediger, H., Balota, D., & Watson, J. (2001). Spreading activation and arousal of false memories. In H. Roediger, J. Nairne, & A. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder. Science conference series* (pp. 95–115). Washington, DC: American Psychological Association.

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*, 803–14.

Sadoski, M., & Paivio, A. (2004). A dual coding theoretical model of reading. In R. B. Ruddell & N. J. Unrau (Eds.), *Theoretical models and processes of reading* (5th ed., pp. 1329–1362). Newark, DE: International Reading Association.

Seamon, J. G., Luo, C. R., Schlegel, S. E., Greene, S. E., & Goldenberg, A. B. (2000). False memory for categorized pictures and words: The category associates procedure for studying memory errors in children and adults. *Journal of Memory and Language, 42*, 120 –146.

Schacter, D. L. (2012). Adaptive constructive processes and the future of memory. *American Psychologist, 67*, 603–613.

Schacter, D. L., Chamberlain, J., Gaesser, B., & Gerlach K. D. (2012) Neuroimaging of true, false, and imaginary memories: Findings and implications. In L. Nadel & W. Sinnott-Armstrong (Eds.), *Memory and law* (pp. 233-262). New York: Oxford University Press.

Schacter, D. L., Israel, L., & Racine, C. (1999). Suppressing false recognition in younger and older adults: the distinctiveness heuristic. *Journal of Memory and Language, 40*, 1–24.

Slotnick, S. D., & Schacter, D. L. (2004). A sensory signature that distinguishes true from false memories. *Nature Neuroscience, 7*, 664-672.

Slotnick, S. D., & Schacter, D. L. (2006). The nature of memory related activity in early visual areas. *Neuropsychologia, 44*, 2874–2886.

Sommers, M. S., & Lewis, B. P. (1999). Who really lives next door: Creating false memories with phonological neighbors. *Journal of Memory and Language, 40*, 83-108.

Stark, C. E. L., Okado, Y., & Loftus, E.F. (2010). Imaging the reconstruction of true and false memories using sensory reactivation and the misinformation paradigms. *Learning and Memory, 17*, 485–488.

Tajika, H., Neumann, E., Hamajima, H., & Iwahara, A. (2005). Eliciting false memories on implicit and explicit memory tests after incidental learning. *Japanese Psychological Research, 47*, 31-39.

Tse, C. S., & Neely, J. H. (2005). Assessing activation without source monitoring in the DRM false memory paradigm. *Journal of Memory and Language, 53*, 532-550.

Tse, C. S., & Neely, J. H. (2007). Semantic and repetition priming effects for Deese/Roediger—McDermott (DRM) critical items and associates produced by DRM and unrelated study lists. *Memory & Cognition, 35*, 1047-1066.

Underwood, B. J. (1965). False recognition produced by implicit verbal responses. *Journal of experimental psychology, 70*, 122.

Van Damme, I., Menten, J., & d'Ydewalle, G. (2010). The effect of articulatory suppression on implicit and explicit false memory in the DRM paradigm. *Memory, 18*, 822-830.

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.

Weinstein, Y., & Shanks, D. R. (2010). Rapid induction of false memory for pictures. *Memory, 18*, 533-542.

Wilkinson, S., & Howe, M. L. (2018). *Persistence pays off: Neutral and emotional false memories serve as more effective problem-solving primes than true memories following a retention interval*. Manuscript submitted for publication.

Appendix 3A: DRM lists

Pen	Bread	Chair	Foot	Fruit	Shirt	Cup	Needle	Car
pencil	butter	table	shoe	apple	blouse	mug	thread	truck
fountain	sandwich	legs	hand	vegetable	sleeves	saucer	pin	bus
leak	rye	seat	toe	orange	pants	tea	eye	train
quill	jam	couch	sandals	kiwi	tie	coaster	point	Jeep
marker	milk	desk	soccer	citrus	button	handle	prick	Ford
Bic	flour	recliner	heel	pear	jacket	coffee	thimble	keys
scribble	dough	sofa	ankle	banana	collar	straw	haystack	garage
crayon	crust	wood	arm	berry	vest	goblet	thumb	tyre

Appendix 3B: Example of degraded representation of the pictures in the perceptual closure task (*car*)

Supplementary materials

Pilot Experiment

Method

Participants

A total of 30 participants (M_{age} = 20.00, *SD* = 2.44) from Maastricht University participated in the experiment in exchange for credit points or a financial reward of €7.50. Seven were males and 23 were females, with age ranging from 18 to 26 years old. All participants were required not be involved in similar memory-related experiments. The experiment was approved by the ethical committee of the Faculty of Psychology and Neuroscience, Maastricht University.

Materials

Study phase. In the current experiment, 14 DRM wordlists that could be represented pictorially were chosen from Peters, Jelicic, and Merckelbach (2015). Each DRM list included 12 associated words (e.g., *hill, valley, climb, summit, Alps*, etc.) and these words were all related to a non-presented target or "critical lure" (i.e., *mountain*). We used pictorial versions of the DRM words (see also Israel & Schacter, 1997; Schacter, Israel, & Racine, 1999). The pictures were collected by means of a Google Images search. For example, for the word "*hill*", we searched for a picture of a "*hill*" using the Google Image search engine. Because of the nature of DRM word associates, around 44.6% (*n*=75) of the pictures referred to simple objects (e.g., *pencil*) and 55.4% (*n* =93) of the pictures related to more complex concepts (e.g., *climb*). The number of pictures that have a background and the number of images showing objects were balanced in different conditions. The pictures were presented digitally for 1000ms, with 1000ms inter-stimulus interval using E-Prime 2.0 software. Pictures within each list were presented in descending associative order.

Recognition test. The recognition test contained 94 pictures. Of these pictures, 42 pictures were presented in the study phase (e.g., pictures of *water, stream, lake*, etc.; three randomly selected pictures per list), 14 were pictures of critical lures (e.g., image of *river*), 14 were not presented but related pictures (e.g., images depicting *waterfall*) and 24 were not presented and unrelated pictures (e.g. pictures of *beard, circus, wire*, etc.). Each picture was shown briefly for 100 ms and there was a fixation cross on screen lasting 1500 ms after each picture.

Perceptual closure task. The perceptual closure task (PCT) included 136 pictures in total. Of all these pictures, 84 were studied pictures, 14 were pictures of critical lures, 14 were new pictures of the critical lures and 24 were new and unrelated pictures. It should be noted that there were two sets of pictures for the critical lures. One set (14 pictures) were shown in the recognition test, the other set (14 new pictures of the critical lures) were not included in the recognition test but were tested directly in PCT (see Figure 3.1). For the 84 studied pictures, half were presented both in the DRM lists and in the recognition test while the other half were only shown in the DRM lists but not included in the recognition test. The pictures were displayed with a distortion filter that enabled the pictures to become clearer over time. We included 9 gradual gradations of blur, and the last one was the original picture. Each blur was presented for 1 second and participants had to respond within 10 seconds. This task recorded both the accuracy and the reaction times for the identification of the pictures using the software E-Prime 2.0. The distorted pictures were created using GIMP 2.0.

Design and Procedure

A 2 (Memory type: studied items vs. critical lures) ×2 (Recognition presentation: yes vs. no) within-subject design was used. In order to control the possible priming effect from the recognition test, a random half of the studied pictures in the PCT were presented in the recognition test and the other half were not shown during the recognition test. For the critical lures in the PCT, two versions of pictures for the same critical lure were prepared; one version was presented in the recognition test and the other was not. Participants were tested individually for approximately 45 minutes in lab facilities. First, participants were shown 14 pictorial DRM lists, with order of the lists randomized. Participants were instructed to memorize each picture and think about the concept of the picture when they saw it on the screen. After the presentation of all pictures, participants played Tetris for 5 min (filler task). Next, participants received the recognition test, during which pictures were subsequently presented and participants were asked to answer “yes” or “no” in a pop-up window to indicate whether they had seen the picture before or not.

Finally, following five minutes of the filler task, participants received the perceptual closure task. They were asked to press the space bar as soon as they recognized the blurred picture. After pressing the button, a window popped up and asked participants to type in what they thought represented the degraded picture. The pictures were successively presented in a pre-determined random order. Input results and reaction times were measured by the E-Prime 2.0 software.

Results and Discussion

The mean recognition rate for all studied pictures was 87.78% (*SD* = 9.99). The mean false recognition rate for critical lures was 28.04% (*SD* = 12.83).Before analyzing the data from the perceptual closure task, we filtered the data based on the following criterion (see also Otgaar et al., 2015). First, for studied pictures, we only included the reaction times for pictures that participants had recognized in the recognition test. This is to make sure they had formed true memories for these pictures. Second, for critical lures, only when participants falsely recognized the critical lures as presented during the recognition test were the reaction times for these pictures included in the analysis. This again, is to make sure they had formed false memories for the critical lure pictures. The above two criteria did not apply to pictures in the "no recognition presentation" condition as those pictures were not included in the recognition test. Third, pictures that were not identified correctly in the PCT were removed. This applied to all studied pictures, critical lures and unrelated items. For instance, if a participant pressed the button indicating identification of a picture (e.g., *mountain*) but then did not write an answer or filled in an incorrect response (e.g., *land*), then this was removed from the analyses. However, synonyms of the answers were accepted. Less than 5% of the accepted answers were synonyms. Of a total of 2940 answers that could be provided, 90.6% (n = 2665) were correct and 9.4% were blank or incorrect (n = 275).

A 2 (Memory type: studied items vs. critical lures) ×2 (Recognition presentation: yes vs. no) repeated measures ANOVA was conducted to examine the effect of visual false memories on the reaction time performance of PCT. Missing data (3 out of 120 cases) were replaced by the average value in the corresponding condition. First of all, no statistical interaction effect was found, $F(1, 29) = 0.004$, $p = .951$, $\eta^2_{partial} = .0001$ (see Figure 3.5). We did find a statistically significant effect of memory type, $F(1, 29) = 17.45$, $p < .001$, $\eta^2_{partial} = .38$., showing that critical lures (presented in recognition test: M=3694, SD =1262; non-presented in recognition test: M=4017, SD =633) were more slowly identified than studied items (presented in recognition test: M=3222, SD =511; non-presented in recognition test: M=3561, SD =511). We also found a significant main effect of recognition test, $F(1, 29) = 5.84$, $p = .02$, $\eta^2_{partial} = .17$,

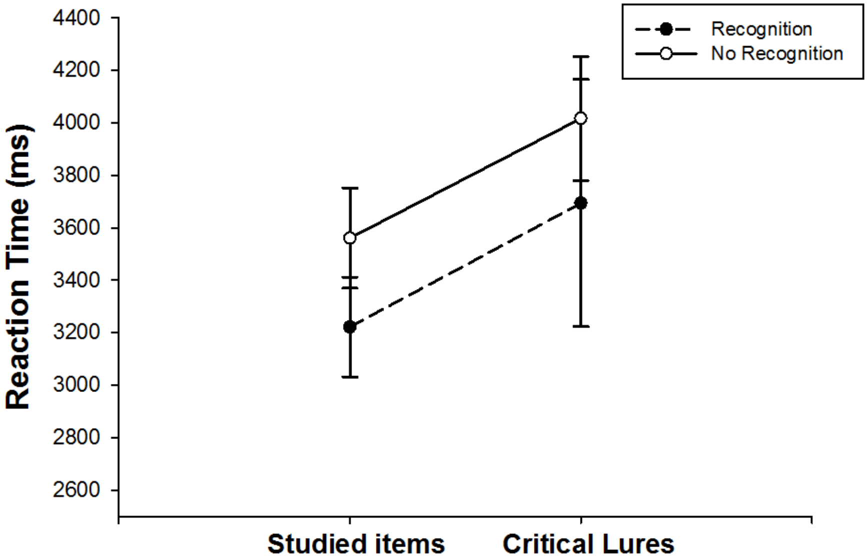


Figure 3.5. Mean reaction times in different Memory type and Recognition presentation conditions. Error bars indicate 95% confidence intervals.

The reaction times for unrelated pictures were also recorded: 83.3% (n = 20) of the unrelated pictures were neither presented during study phase nor shown during the recognition test. The mean reaction time for unrelated pictures was 4547 ms (SD = 681). A series of paired-samples t tests were conducted to compare the four conditions (studied items with or without recognition presentation, critical lures with or without recognition presentation) with unrelated pictures. Results showed unrelated items were identified more slowly than any of the four conditions (ps <0.001).

Exploratory analysis

When we looked into the answers participants provided during PCT, we found that answers to some pictures with various backgrounds were quite divergent. Pictures with backgrounds might introduce alternative interpretations for them. For instance, participants wrote “river” or “landscape” for a picture of “valley”. Those pictures might look ambiguous to participants and hence, took more time to identify. In order to rule out ambiguity as a confounding variable, we deleted fourteen pictures that seemed ambiguous. We re-analyzed the data and the results showed the same pattern in Figure 3.5: no interaction effect was found between memory type and recognition presentation, $F(1, 29)$ =

0.68, p = .679, $\eta^2_{partial}$ = .023; memory type had a main effect on the PCT reaction times, with critical lures identified slower than studied items, F(1, 29) = 26.33, p < .001, $\eta^2_{partial}$ = .48; recognition presentation also had a main effect, F(1, 29) = 9.18, p = .005, $\eta^2_{partial}$ = .24.

Nonetheless, our results could be affected by the type of stimuli that we used. As has been mentioned earlier, in the present experiment, the pictures included not only objects but also pictures depicting a movement or a scene, which resulted that the layout of some pictures were more complex than others. Indeed, research has shown that false memory rates differ when heterogeneous backgrounds or homogeneous backgrounds are used. Howe (2008), for example, presented children with photographs that shared a single background (homogeneous) and photographs with a unique background for each list item (heterogeneous), and found that homogeneous photographs resulted in more false memories. Because some pictures were more complex and had various backgrounds, it is likely that it took more time to process and identify these pictures than simple pictures of a single object thereby possibly impacting our findings. Although our exploratory analysis did not support this possibility (the results stayed unaffected after deleting ambiguous pictures), a second experiment with stricter control is needed. To address this issue, we presented participants with only pictures of simple objects (e.g., pen, bread) at a single white background in the official experiments. In this way, pictures only varied between each other in terms of the objects that they represented.

Part II

Consequences of Belief and Recollection

CHAPTER 4

Undermining Belief Changes Problem Solving

This Chapter is an adapted version of the following paper:

Wang, J., Otgaar, H., Howe, M. L., Smeets, T., & Merckelbach, H. (2017). Undermining belief in false memories leads to less efficient problem-solving. *Memory*, 25, 910-921.

Abstract

Memories of events for which the belief in the occurrence of those events is undermined, but recollection is retained, are called nonbelieved memories (NBMs). The present experiments examined the effects of NBMs on subsequent problem-solving behavior. In Experiment 1, we challenged participants' beliefs in their memories and examined whether NBMs affected subsequent solution rates on insight-based problems. True and false memories were elicited using the Deese/Roediger-McDermott (DRM) paradigm. Then participants' belief in true and false memories was challenged by telling them the item had not been presented. We found that when the challenge led to undermining belief in false memories, fewer problems were solved than when belief was not challenged. In Experiment 2, a similar procedure was used except that some participants solved the problems one week rather than immediately after the feedback. Again, our results showed that undermining belief in false memories resulted in lower problem solution rates. These findings suggest that for false memories, belief is an important agent in whether memories serve as effective primes for immediate and delayed problem-solving.

Keywords: Nonbelieved memory; Belief; Recollection; False memory; Problem-solving

Jack has a memory in which he put his hand in a cage at the Philadelphia Zoo and his left wrist was bitten by a monkey. Years later, his mother assured him that it never happened. He does not believe the horrible event actually happened, but he cannot stop having vivid 'recollections' or 'memories' concerning the event (http://www.falsememoryarchive.com/). The question is, when Jack comes across monkeys in the zoo, will he stay away from them?

Memories of events for which the belief in the occurrence of those events has been undermined but the recollection has been preserved, are called nonbelieved memories (NBMs). This recently studied phenomenon turns out not to be rare, with more than 20% of people reporting that they have vivid but non-believed autobiographical memories (Mazzoni, Scoboria, & Harvey, 2010). A lingering question is whether these NBMs have any impact on behavior (e.g., avoiding monkeys). The current experiments delve into this question by examining the (in)dependent behavioral consequences of beliefs and recollections on performance on subsequent problem-solving tasks.

Previous research on the behavioral consequences of memories has predominantly focused on *believed* memories (Scoboria, Jackson, Talarico, Hanczakowski, Wysman, & Mazzoni, 2014), with few studies looking at the behavioral consequences of *nonbelieved* memories. An important reason why we focus on the behavioral consequences of NBMs is that there is a new line of research that has demonstrated that belief and recollection are independent constructs that can have differential effects on behavior (Scoboria, Mazzoni, Kirsch, & Relyea, 2004; Scoboria, Talarico, & Pascal, 2015). Here, *belief* refers to the truth-value related to the occurrence of an event, whether or not a recollection is present. *Recollection* refers to the mental re-experiencing of an event (e.g., Rubin, 2006). Various theorists argue that memories contain key components that lead to both a sense of re-experiencing the event and a belief that the event actually occurred (e.g. James, 1890/1950; Brewer, 1996; Schacter, 1996; Tulving, 1985).

In recent years, this view has gained more attention, with the distinction between belief and recollection being supported by empirical research. For example, Scoboria et al. (2014) used structural equation modelling and found that factors that predicted recollection (e.g., perception, re-experiencing) were distinct from factors that predicted belief (e.g., plausibility), suggesting a dissociation between recollection and belief. For most of our memories, belief and recollection both contribute to remembering. Scoboria and Talarico

(2013) found that for believed autobiographical memories, belief and recollection ratings both tend to be at the high end (above 7) on a 1 to 8 likert scale.

However, in other cases, only belief or recollection is present. For example, there are family stories (e.g., your birth) that one believes occurred but cannot recollect. There are also NBMs where vivid recollections of events exist (e.g., believing you actually saw Santa Claus putting presents under the tree as a child) but beliefs for these events are undermined (e.g., by acquiring knowledge that Santa Claus is a fictional character) (Mazzoni, Scoboria, & Harvey, 2010; Otgaar, Scoboria, & Mazzoni, 2014).

An examination of NBMs may help uncover how belief and recollection interactively lead to behavioral outcomes. For instance, in studies investigating how believed memories impact behavior, participants exhibited superior public speaking performance and higher levels of exercise if believed memories of relevant positive experiences were activated (see Biondolillo & Pillemer, 2015; Pezdek & Salim, 2011). However, because these 'memories' were both believed and recollected, it is hard to know whether it was recollection, belief, or both that were responsible for the changes in subsequent behavior.

Research concerning false memories has also examined the impact of belief on behavior. For example, Bernstein and Loftus (2009) reviewed a number of studies where researchers created false memories about childhood events, such as being ill after eating egg-salad. These memories resulted in a subsequent reduction in eating egg-salad. Again, however, it is hard to determine the source of change in subsequent behavior as belief and recollection were confounded. That is, participants developed a false belief about the false memory (event) with approximately a quarter of participants reporting having recollections of the false event (also see, Scoboria, Mazzoni, Jarry, & Bernstein, 2012). Indeed, these authors were acutely aware of this problem as Laney, Morris, Bernstein, Wakefield, and Loftus (2008, p. 291) noted that, "the data in the present paper represent some false memories and some false beliefs. But because it is awkward to say 'false memories and false beliefs' repeatedly, we generally just use one term (either "false memory" or "false belief") to encompass the notion of planting a false entity."

To our knowledge, there is no research that has directly and experimentally tested whether it is false belief or false recollection that affects behavior (but see Otgaar, Moldoveanu, Wang, & Howe, 2016). Recently, Bernstein, Scoboria, and Arnold (2015) conducted a mega-analysis on previously published food-preference experiments (see above) and concluded that for false events, belief is more important than memory in modifying food preference. Indeed, they stated that, "[c]ompared to memory of past events, belief in the occurrence of past events is more important for altering attitudes and behaviors" (p. 6). However, in all the experiments reviewed, belief and recollection for the suggested false events were neither intentionally nor clearly manipulated separately. Moreover, NBMs were not addressed in any of the experiments. Perhaps more importantly, food preferences may result in part from decision-making processes that are analytic and occur consciously. For instance, participants might reason: "Since egg-salad made me ill, I'd better not eat it." However, in a problem-solving process that involves intuitive thinking or 'Aha!' experiences (e.g., insight-based problem-solving; Bowden, Jung-Beeman, Fleck, & Kounios, 2005; Howe, Garner, Dewhurst, & Ball, 2010), there is usually no explicit reasoning about facts/knowledge. Recollection might play a vital role independent of belief in insight-based problem-solving behavior where people are unaware of the processes underlying the solutions to these types of problems.

The Current Experiments

The main purpose of the current experiments is to examine the impact of nonbelieved memories on insight-based problem-solving behavior. We decided not to use the food preference paradigm to elicit false beliefs/memories because this paradigm only allows one false belief or false memory to be created per participant. Because so few false beliefs and memories are created, it is even more difficult to produce the necessary number of nonbelieved memories. Therefore, we opted for a method that leads to high and reliable levels of false memories, the Deese/Roediger-McDermott (DRM) paradigm (Deese, 1959; Roediger & McDermott, 1995). In this paradigm, participants are presented with lists of associated words (e.g., *butter, food, eat, sandwich*) that are all related to a non-presented critical lure (i.e., *bread*). Participants not only correctly remember (recall, recognize) items presented on the lists but also form false memories for critical lures that were not presented. In the present experiments, after false memories were formed with the DRM paradigm, participants were challenged on their responses by telling them certain items were not presented in an attempt to create nonbelieved false memories. Data from our lab (Otgaar et al., 2016) have confirmed that nonbelieved false memories can be created using the DRM paradigm. Also by using this paradigm we were able to examine the impact of beliefs on true memory and explore the effects of nonbelieved true memories on behavior.

Following the challenges to true and false memories, participants were asked to solve compound remote associate task (CRAT) problems, in which solutions referred to nonbelieved and believed words (true and false memories from DRM lists). A CRAT problem consists of three words (e.g., *Board/Mail/Magic*). To solve the problem, participants have to come up with a word that could link all the three words (in the example given above, the answer was *Black*). Howe et al., (2010) presented participants with DRM lists and then asked participants to solve CRAT problems whose solutions were critical lures for the DRM lists. They found that CRAT problems primed by false memories for critical lures were solved more frequently and significantly faster than problems that were not primed. Subsequent research typically showed the priming effect of believed false memories on CRAT problems (Howe, Garner, Charlesworth, & Knott, 2011; Howe, Wilkinson, Garner, & Ball, 2016); and found that the priming effect of false memories was similar to or even stronger than that observed for true memories (Howe, Threadgold, Norbury, Garner, & Ball, 2013; Howe, Wilkinson, Monaghan, Ball, & Garner, 2013). Based on Bernstein et al.'s work (2015), we predicted that if belief is more influential than recollection when it comes to impacting subsequent behaviors that ostensibly require non-conscious problem-solving processes, no priming effect would be found after beliefs for false memories are withdrawn. However, if recollection plays a more vital role in priming the CRATs than beliefs, then nonbelieved false memories should prime as many CRATs as believed false memories.

We were also interested in exploring individual differences in the formation of NBMs. Social feedback has been found to be one of the main contributors to fostering NBMs (Scoboria, Boucher, & Mazzoni, 2015) and hence we included the Gudjonsson Compliance Scale (Gudjonsson, 1989) in order to examine individual differences in social compliance. Furthermore, a scale measuring dissociative symptoms was administered because dissociation has frequently been linked to the formation of false memories (Giesbrecht, Lynn, Lilienfeld, & Merckelbach, 2008).

Experiment 1

Method

Participants

Before the recruitment of participants, we ran a power analysis using G*Power 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007), with an estimated power of 0.80 and a medium effect size of 0.25 (*f*). The power analysis revealed that 34 participants needed to be tested. A total of 36 students from Maastricht University participated in the experiment in exchange for credit points or a financial reward of €7.50. Two participants were excluded because they did not complete the CRAT problem-solving session, thus leaving 34 participants (14 males and 20 females). All participants were native English speakers, aged between 17 and 34 (M_{age} = 21.6, *SD* = 3.26). The experiment was approved by the ethical committee of the Faculty of Psychology and Neuroscience, Maastricht University.

Materials

DRM lists. Sixteen DRM lists were used in our experiment. These lists have successfully been used in previous research (e.g., Howe, Garner, & Patel, 2013). Each DRM list included 12 associated words (e.g., *butter, food, eat, sandwich*) and these words are all related to a non-presented target or "critical lure" (i.e., *bread*). Importantly, to eliminate possible item effects arising from differences between a studied item and a critical lure, for eight lists the first list word was replaced by the critical lure (see also Howe et al., 2013). Thus, these "critical lures" are no longer "false" memories as they now become "true" studied items presented as part of the list. The other eight lists were standard DRM lists that had the corresponding eight critical lures. The recognition task contained 56 words, of which 24 items were presented items from the DRM lists, 8 were non-presented critical lures from the DRM lists, and 24 words were not presented and served as unrelated lures.

CRAT problems. We used 24 CRAT problems in this experiment (taken from Howe et al., 2013). Each CRAT was comprised of three words (e.g., *crust, stale, French*), all of which could be solved by a single linking word (i.e., *bread*). Sixteen CRATs were primed by the preceding sixteen DRM lists: half of the CRAT problems were primed by lists whose false memories (critical lures) were the solution words and the other half were primed by lists whose true memories (studied items) were the solution words (see Appendix). The other eight CRATs were not primed and served as an unprimed control condition. The mean solution rate and solution time for each CRAT were known from previous research (Howe et al., 2013). Figure 4.1 illustrates the alignment of DRM items and CRAT problems.

Gudjonsson Compliance Scale (GCS; Gudjonsson, 1989). The GCS is a self-report questionnaire measuring the degree of compliance. It contains 20 true/false statements (e.g., “I often give in to people when I am under pressure”). The total score of GCS ranges from 0 to 20, with higher scores indicating more compliant tendencies. The CGS has an internal consistency of .71 and a test-retest reliability coefficient of .88.

Dissociative Experiences Scale (DES; Bernstein & Putman, 1986). The DES measures the degree to which people experience dissociative symptoms. It consists of 28 items (e.g., “Some people find that sometimes they are listening to someone talk and they suddenly realize that they did not hear part or all of what was said.”) and participants have to select what percentage of time this happens to them from 0% to 100% with 10% increments. It has a good internal consistency, with Cronbach α= .92.

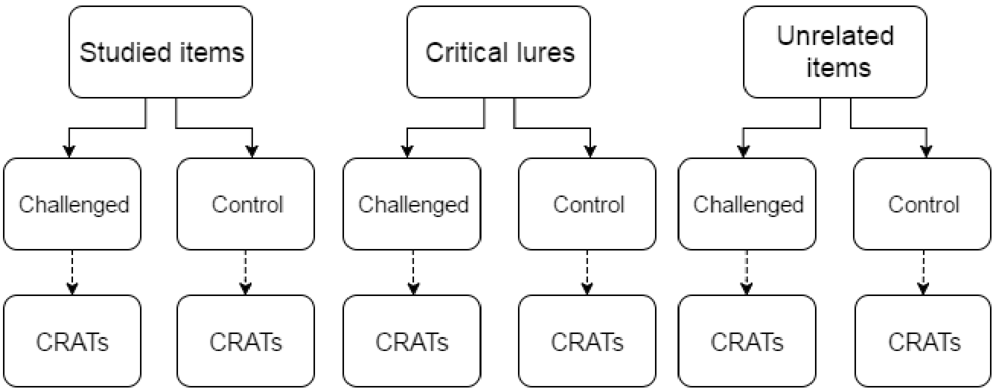


Figure 4.1. Diagram of alignment between DRM items and CRAT problems.

Design and Procedure

A 3(Memory Type: critical lures, studied items, unrelated items) × 2 (Belief: challenged vs. control) within-subject design was used. Participants were tested individually for approximately 50 minutes in lab facilities at the faculty. There were four phases.

1) Study phase: DRM words were presented to participants on a computer screen. Participants were instructed to remember as many words as they could. Each word was visually presented for 1500ms, with 500ms inter-stimulus interval using the program Visual Basic. The sequence of the words within a list was fixed, but the order of the lists was randomized.

2) Recognition phase: After a distractor task (playing the game Bejeweled for 3 min), participants were involved in a recognition task, in which 8 critical lures, 24 studied words, and 24 unrelated words were included to examine false memories and true memories, respectively. Participants were asked to identify the word on the screen as to whether it had been presented by clicking on a “Yes” or “No” button. Next, participants completed the dissociation questionnaire (DES) before they moved to the challenging phase (see below).

3) Challenging phase: Participants were told that the computer graded their answers and gave them feedback on their performance on the previous recognition test. Before their responses were challenged, participants were told why our memories were sometimes unreliable and they were shown an extra DRM list to illustrate how a DRM list could lead to the formation of false memories. This explanation was given so participants understood why feedback was presented.

In the challenged belief condition, when a certain target word (e.g., *bread*) appeared on the screen, a label beneath the word popped up stating that “Sorry, your previous answer was incorrect. This word was not presented.” In the control condition, the feedback was “Congratulations, your answer was correct. This word was presented.” In this way, we attempted to create nonbelieved and believed memories, respectively. When the feedback was provided, the experimenter showed participants a printed fake proof where presented words were listed and gave oral social feedback asking them to rethink their previous answer. Then immediately after the challenge for each word, participants rated their memory and belief for that word on 1-8 Likert scales (i.e., “Do you have a memory for this word?”, 1= no memory at all, 8 = clear and complete memory; “Do you believe that this word was presented to you?”, 1= definitely did not happen, 8 = definitely did happen; adapted from Scoboria et al., 2004). The experimenter explained thoroughly to the participant the difference between memory (i.e., recollection) and belief. The sequence of all challenged words was randomized. In total, 56 words from the recognition test were given feedback, but we were only interested in the 24 target words that served as solutions to the corresponding CRAT problems. The belief in half of the target words (4 critical lures, 4 studied words, and 4 unrelated) was always challenged regardless of their original recognition responses. The other words were in the control condition in which participants' belief in their presence was not challenged.

4) Problem-solving phase: Twenty-four CRAT problems were presented in which the correct answers were the target words in the challenging phase. The twenty-four CRATs were assigned to six (3×2) conditions. An ANOVA revealed no statistically significant differences between mean solution rates across different item types ($F(2, 18) = 0.01$; $p = 0.99$) and different belief conditions ($F(1, 18) = 0.001$; $p = 0.97$). There were no solution time differences across item types ($F(2, 18) = 0.02$; $p = 0.98$) and belief conditions ($F(1, 18) = 0.16$; $p = 0.70$). These analyses were done as a manipulation check to make sure there was no baseline difference in reaction times and solution rates among clusters of CRATs across conditions.

Participants were falsely told that the problem-solving phase was a separate experiment aimed at examining how personality (i.e., the dissociation questionnaire) affected problem-solving style. Participants were instructed that three words would be presented on the screen and their task was to come up with a word that could link all the three words. Participants were given an example first (e.g., the answer to the problem *apple/family/house* was *tree*), followed by one practice CRAT problem that they had completed themselves before they began the test CRATs. Problems were presented in a random order. A countdown timer appeared in the upper right corner of the screen and participants were asked to type their solution within 60 seconds. Upon completion of each CRAT, no correct answer was given to lower the risk that participants would connect the memory task with the problem-solving task. Solution rates and times were recorded by the computer. After all sessions, participants filled in the compliance scale (GCS) and were debriefed about the purpose of the study.

Results and Discussion

Recognition rates

The mean recognition rate for all studied items was 74.50% ($N = 608$). The mean false recognition rate for critical lures was 69.13% ($N = 188$), which is consistent with previous research (e.g., Blair, Lenton, & Hastie, 2002). Participants falsely recognized 14.63% (N = 20) of non-presented unrelated items.

The mean recognition rate in each condition is shown in Table 4.1. Recognition rates were analyzed using a 3(Memory Type: critical lures, studied items, unrelated items) × 2(Belief: challenged vs. control) repeated measures ANOVA. No interaction was found, $F(2, 66) = 1.57$, $p = .22$, partial $\eta^2 = .05$. There was a main effect of Memory Type, $F(2, 66) = 185.51$, $p < .001$, partial $\eta^2 = .85$, where Bonferroni post-hoc analyses showed that participants recognized statistically more critical lures and studied words than non-presented words ($ps < .001$). There was no main effect of Belief, $F(1, 33) = 0.91$, $p = .35$, partial $\eta^2 = .03$, indicating that there were an equivalent number of true and false memories in both the challenged belief and control conditions.

Table 4.1

Recognition rates in different Memory Type and Belief conditions (M, 95%CI)

	Memory Type		
	Critical lures	Studied items	Unrelated items
Challenged belief	0.68 [0.57, 0.80]	0.88 [0.82, 0.93]	0.10 [0.06, 0.15]
Control condition	0.70 [0.60, 0.80]	0.85 [0.77, 0.92]	0.19 [0.11, 0.27]

Nonbelieved and believed memories

We recorded memory and belief ratings for each word after the challenge manipulation. Analogous to previous research, we employed the following criteria for nonbelieved memories: recollection needed to be rated at least 2 scale points higher than belief (see Mazzoni, Clark, & Nash, 2014; Otgaar, Moldoveanu, Wang, & Howe, 2016), and within this criterion, the recollection rating should be at least 3. For believed memories, we set the same criterion for recollection rating (at least 3) as in nonbelieved memories, and belief rating should be equal to, or above, 3. Within the categories of nonbelieved and believed memories, if the item was a critical lure, it was a false memory; if it was a studied item, it was a true memory. For unrelated items, there were two categories: items with no belief and no memory (ratings ≤ 2) and items with no memory but belief. Table 4.2 shows the mean Memory and Belief ratings for words in each condition.

Table 4.2

Mean memory and belief ratings in each kind of induced memory (CI: confidence interval; n: number of participants contributing to the mean score; N: number of items contributing to the mean score)

	Memory Rating (95% CI)	Belief Rating (95%CI)	Memory-Belief difference
Nonbelieved False Memory	5.02 [4.61, 5.44] (*n*=27; *N*=76)	1.42 [1.12, 1.71] (*n*=27; *N*=76)	3.60
Nonbelieved True Memory	5.75 [5.33, 6.17] (*n*=33; *N*=107)	1.96 [1.64, 2.28] (*n*=33; *N*=107)	3.79
No Belief No Memory	1.27 [1.13, 1.40] (*n*=33; *N*=121)	1.23 [1.09, 1.38] (*n*=33; *N*=121)	0.04
Believed False Memory	6.46 [5.97, 6.94] (*n*=34; *N*=124)	7.25 [6.92, 7.59] (*n*=34; *N*=124)	-0.79
Believed True Memory	7.09 [6.81, 7.36] (*n*=34; *N*=122)	7.59 [7.41, 7.78] (*n*=34; *N*=122)	-0.50
Belief with No Memory	1.94 [1.62, 2.26] (*n*=30; *N*=93)	6.36 [5.85, 6.86] (*n*=30; *N*=93)	-4.42

After the challenge manipulation, 97.1% of the participants (n = 33) had formed at least one nonbelieved true memory for studied items, with an average number of 3.15 (SD = 1.05) nonbelieved memories. 79.4% of the participants (n= 27) had developed at least one nonbelieved false memory, with an average number of 2.24 (SD = 1.52). In the challenged belief condition, participants formed nonbelieved false memories for 55.88% of the critical lures and formed nonbelieved true memories for 78.68% of the studied words. In the control condition, participants formed believed false memories for 91.18% of the critical items and formed believed true memories for 89.71% of the studied words.

Solution rates of CRATs

The mean CRAT solution rates (in proportions) were calculated for each participant. We focused on words that were effectively manipulated into believed/nonbelieved memories. We labelled the effect of successfully challenging or lowering participants' beliefs as "undermining." Seven participants had formed either zero nonbelieved false memories or zero nonbelieved true memories, hence the CRAT solution rates of these cases were treated as missing values. However, these participants had CRAT solution values in the believed memory conditions, thus we used a multiple imputation method (Schafer, 1997) to impute missing data values. In total, 3.9% of the data (8 out of 204 cells) were imputed over five cycles of imputations. We compared the CRAT solution rates in the undermining belief condition with the CRAT solution rate in the control condition for the following three memory types: critical lures, studied items, and unrelated items[5]. Interestingly, undermining belief led to different results for false and true memories. In the critical lures condition, undermining belief (M = 0.36, SD = 0.29) led to statistically lower solution rates than control (M = 0.49, SD = 0.28), t = -2.04, df_{pooled} = 884, p = .04. However, for studied items, undermining belief increased solution rates significantly (undermining belief, M = 0.52, SD = 0.27; control condition, M = 0.40, SD = 0.25, t = 2.04, df_{pooled} = 3353, p =.04; see Figure 4.2). We compared the CRAT solution rates in the control condition to see whether the results in the control condition were consistent with previous research. Paired samples t-tests showed that false memories did not differ significantly from true memories in priming the CRATs (p =.27); false and true memories both primed more CRATs than unrelated items (p = .008; p = .02). Thus, the results in the control condition replicate previous findings on the consequences of false memories on problem-solving (e.g., Howe et al., 2013).

[5] In this analysis, words were evaluated as nonbelieved memories when memory ratings for them were two points higher than belief ratings. When we adopted a stricter criterion (i.e., memory ratings are at least three points higher than belief ratings), we found a same interactive pattern and the true memory effect was less pronounced (see supplementary document).

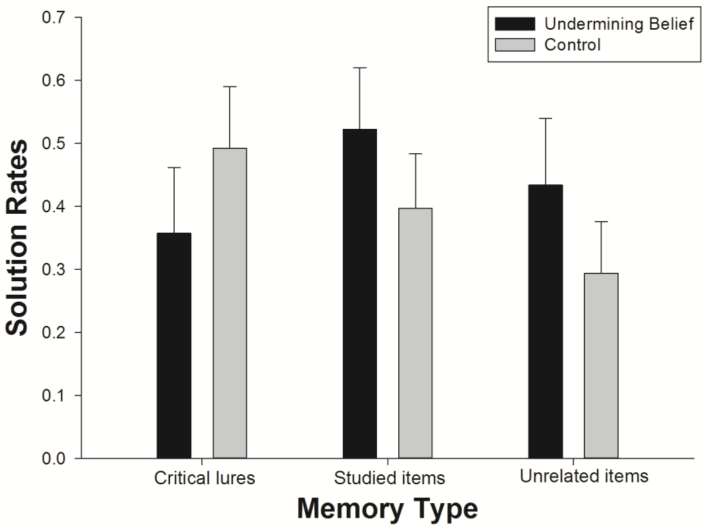


Figure 4.2. Mean solution rates in different Memory Type and Belief conditions (Experiment 1). Error bars indicate 95% confidence intervals.

As has been done in previous related work (e.g., Otgaar, Howe, van Beers, van Hoof, Bronzwaer, & Smeets, 2015), we also performed an additional analysis by focusing only on the items that participants recognized as "presented" in the recognition test. When words with "yes (presented)" recognition responses only were included in analysis, a similar interaction effect between memory and belief was found, $F(2, 64) = 5.36$, $p < .01$, partial $\eta^2 = .14$. No main effect of belief was found, $F(1, 32) = 0.35$, $p = .56$. No main effect of memory was found, $F(2, 64) = 2.37$, $p = .10$.

Solution times of CRATs

Mean solution times of CRATs (in seconds) in each condition were calculated. We were particularly interested in comparing the solution times between nonbelieved and believed memories. In some cases, participants solved no CRATs under the priming of (non)believed memories, thus solution times in that condition were counted as missing. In total, there were 19.11% ($N = 39$) of the cases where solution times were missing. We conducted several paired sample *t*-tests and again found decreasing belief had different effects for false and true memories. In the critical lures condition, nonbelieved false memory (M = 17.75, SD = 9.07) primed problems as fast as believed false memory (M = 14.47, SD = 7.16; p = .27); however, in the studied items condition, undermining belief (M = 12.80, SD = 5.18) resulted in faster solution times to CRAT problems than in the control condition (M = 20.33, SD = 10.30; p = .003). Undermining belief (M = 15.33, SD = 9.58) and control conditions (M = 18.28, SD = 12.15) did not differ in solution times for unrelated items (p = .77). [6]

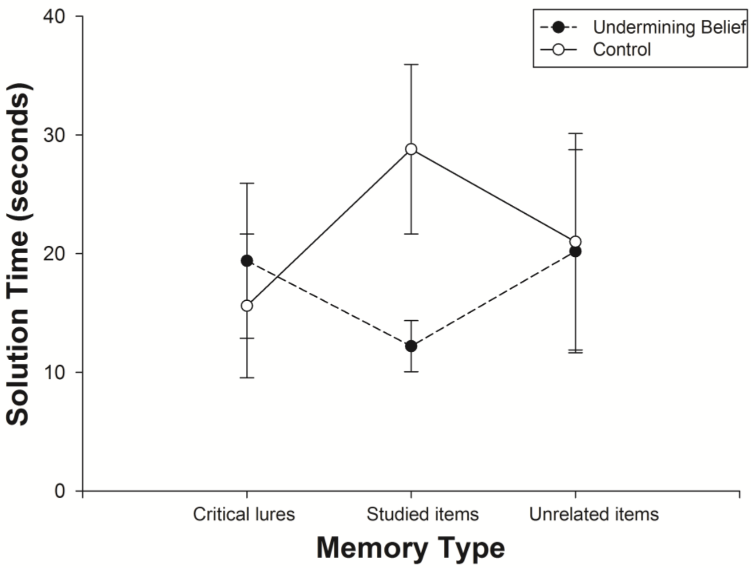


Figure 4.3. Mean solution times in different Memory Type and Belief conditions (Experiment 1). Error bars indicate 95% confidence intervals.

Exploratory Analysis

We conducted Pearson's correlations between the scores on the GCS and the different memories (believed true/false memory; nonbelieved true/false memory). We found a statistically significant correlation between GCS scores and false recognition rates (r (32)

[6] The pattern of results was basically similar when a three-point criterion (i.e. memory recollection was rated at least 3 points higher than belief) for nonbelieved memories was adopted.

= .38, p = .03): the higher scores on the GCS, the higher false recognition rates. No correlation between GCS scores and true recognition rates was detected. As GCS scores measure compliance, we expected that the more compliance participants exhibited, the more nonbelieved memories they would report. However, no statistically significant correlation was found between GCS scores and number of nonbelieved false memories (r (32) = .31, p = .07).

Correlations between DES scores and different kinds of memories were also analyzed. We found a statistically significant negative correlation between dissociative symptoms and true recognition rates (r (32) = -.44, p = .01). Thus, participants scoring high on dissociation had worse memories for presented words. No correlation was found between DES scores and false recognition rates. Interestingly, dissociation scores correlated negatively with the number of nonbelieved unrelated items (r (32) = -.36, p = .04) but not nonbelieved true memories (r (32) = -.08, p = .64) or nonbelieved false memories (r (32) = -.01, p = .94). Of course, we had a small sample size for executing correlation analyses, so the results of these analyses should be cautiously treated.

Experiment 1 examined the behavioral consequences of belief and recollection. We found that in the control condition, false memories primed the CRATs as efficiently as true memories, a finding consistent with previous research (Howe et al., 2010). Interestingly and for the first time, our data also showed that, without belief, the priming effect of false memories was changed under the conditions tested. This is in line with research by Bernstein and colleagues (2015). That is, when belief in false memories was withdrawn, participants solved fewer CRAT problems. This result constitutes the first experimental attempt that shows that nonbelieved (false) memories impact problem-solving behavior and do so differently than true memories.

Some might argue that believed false memories primed more CRATs than nonbelieved false memories because a higher number of believed false memories were created. However, our data on the other memory types did not support this idea. In the studied item condition, there were more believed true memories induced, but believed true memories did not prime more CRAT problems than nonbelieved true memories. Others might argue that memory for an event is different from memory for a word and the implication of studying NBMs for words might not be so illuminating. Indeed, an event consists of multiple elements and usually a recollection of an event contains more vivid details than a recollection of a word.

Hence, nonbelieved memories for events might exhibit greater behavioral impact than nonbelieved memories for words.

Although our data suggest that for false memory, belief might play an active role in problem solving, our study is still preliminary and needs replication. Furthermore, the results of this experiment were somewhat limited because of the following. First, in this experiment, there were only four CRATs in each condition. When we attempted to induce nonbelieved memories for the four solution words to these CRATs, not all of the items could be successfully transformed into nonbelieved memories. The average nonbelieved memory rate of critical lures and studied items was in the 55%-80% range. If there are more items in each condition, more NBMs can be created and thus, the effects of beliefs and recollections can be better investigated. Second, the CRAT problems in each condition were fixed; that is, they not completely randomized for every participant. Although the mean solution rates and times for CRATs were counterbalanced across conditions, it is unknown whether the difficulty of CRATs impacted our results. In order to address these issues and thus, replicate our results, we conducted an additional experiment.

In Experiment 2, we assigned more CRAT problems to each condition and CRAT problems were no longer fixed in each condition. Furthermore, having established the immediate effect of undermining belief on problem solving, Experiment 2 explored the long-lasting effects of nonbelieved memories on problem-solving behavior. From a theoretical perspective, this is important because previous studies have found that the superior priming effect of false memories emerged particularly after a 1-week delay while the priming effect of true memories declined (Howe et al., 2013). One possible explanation is that true memories decay faster than false memories because true memories are often *other-generated* (e.g., presented on a list by the experimenter), whereas false memories tend to be *self-generated* [i.e., occurring spontaneously and automatically as a result of internal semantic activation (Howe, Garner, Threadgold, & Ball, 2015)]. This pattern corresponds to findings from previous studies on false memories and food preferences in which the behavioral impact of belief could last for months (Bernstein & Loftus, 2009). Furthermore, previous research has shown that false beliefs can last up to four months (see Geraerts, Bernstein, Merckelbach, Linders, Raymaekers, & Loftus, 2008; Laney, Fowler, Nelson, Bernstein, & Loftus, 2008) and experimentally evoked nonbelieved false memories endure for as long as a month (Otgaar, Scoboria, & Smeets, 2013).

Because the behavioral effects of false beliefs may be long-lasting and because belief is more easily manipulated than memory, one could anticipate the following. If belief is the more active agent in guiding behavior than memory, then undermining belief should lead to behavioral effects even after a delay. As belief is assumed to be part of knowledge related to the self (e.g., Scoboria et al., 2004), the enduring behavioral effects of belief would be especially evident for false memories as they are the result of internal associative activation and thus, self-generated. Hence, in Experiment 2, half of the participants had to complete the CRAT problem-solving task immediately and the other half following a one-week delay.

Experiment 2

Method

Participants

A power analysis indicated 70 participants were needed when a power of 0.80 and a medium effect size of 0.27 (*f*) were estimated. A total of 71 participants were tested in exchange for credit points or a financial reward of €7.50. The sample consisted of 22 males and 49 females, with a mean age of 22.3 years old (*SD* = 5.97). 84.5% of the participants (*n* = 60) were native English speakers, and 15.5% (*n* = 11) of the participants were fluent in English, but used English as a second language.

Materials

Twenty-four DRM lists were used in Experiment 2. Each list contained 10 associates. The recognition phase included 12 non-presented critical lures from 12 of the DRM lists, 48 studied items (in which 12 were targeted items for belief manipulation), and 36 unrelated items. Twenty-four CRATs whose answers were 12 critical lures and 12 studied items (corresponding to the 24 DRM lists) were used. The GCS and DES questionnaires were administered to participants as well (see Experiment 1).

Design and Procedure

A 2 (Time Interval: immediate vs. 1 week) × 2 (Memory Type: critical lures vs. studied items) × 2 (Belief: challenged vs. control) mixed design was used, where the first factor was between-subjects and the other two were within-subject factors. Thirty-six participants were randomly allocated to the immediate condition and 35 to the 1-week delay condition. Experiment 2 followed the same procedure as Experiment 1, except that 35 of the participants did not finish the problem-solving phase immediately after the challenging phase, but instead, did so one week later. Because the items in the challenged belief and control belief conditions were fixed, we switched the items in these two conditions for half of the participants in each Time Interval group. That is, belief for the same 12 target words was undermined in around half of the participants (*n*= 31), but belief for these words was not challenged in the other participants.

Results and Discussion

Recognition rates

The mean recognition rate for unrelated items was 14.5%. For the targeted studied items and critical lures (i.e., those that served as the solutions to the subsequent 24 CRATs), the mean recognition rates in each condition are shown in Table 4.3.

Table 4.3

Mean recognition rates for critical lures and studied items in different conditions (N: number of recognized items)

		Critical lures	Studied items
Immediate	Challenged belief	69.0% (*N* = 149)	85.6% (*N* = 185)
	Control condition	68.5% (*N* = 148)	81.5% (*N* = 176)
1-week delay	Challenged belief	58.6% (*N* = 123)	81.4% (*N* = 171)
	Control condition	62.9% (*N* = 132)	76.2% (*N* = 160)

Nonbelieved and believed memories

Memory and belief ratings for each word were recorded after the belief manipulation for that word. The same criteria for nonbelieved and believed memories used in Experiment 1 were used here. Table 4.4 shows the mean percentages of critical lures and studied items that were nonbelieved and believed memories.

Table 4.4

Percentages of critical lures and studied items that were nonbelieved and believed memories in each condition

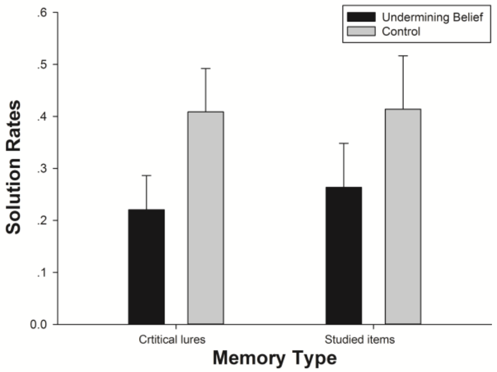
		Critical lures	Studied items
Immediate	Challenged belief (Nonbelieved memories)	58.33% (*N*=126)	61.17% (*N*=132)
	Control condition (Believed memories)	91.67% (*N*=198)	95.33% (*N*=206)
1-week delay	Challenged belief (Nonbelieved memories)	45.23% (*N*=95)	48.09% (*N*=101)
	Control condition (Believed memories)	84.28% (*N*=177)	87.62% (*N*=184)

Solution rates of CRATs

Again, we labelled the effect of successfully challenging or lowering participants' beliefs as "undermining." A 2 (Time Interval: immediate vs. 1 week) × 2 (Memory Type: critical lures vs. studied items) × 2 (Belief: undermining vs control) repeated measures ANOVA was conducted, with Time Interval as a between-subjects variable. Note that there were 14 participants in total who did not form a nonbelieved true or a nonbelieved false memory. These participants' CRAT data were not entered into the analysis. As a result, 34 participants were in the immediate condition and 23 were in the delay condition. There was no statistically significant three-way interaction effect for Time Interval × Memory Type × Belief, $F(1, 55) = 0.14$, $p = .71$, and also no statistically significant two-way interactions. There was a main effect for Belief, $F(1, 55) = 20.68$, $p < .001$, partial $\eta^2 = .27$, which, like Experiment 1, showed that undermining belief resulted in lower solution rates. There was no main effect for Memory Type, $F(1, 55) = 1.37$, $p = .25$, partial $\eta^2 = .02$. Neither was there a main effect for Time Interval, $F(1, 55) = 0.44$, $p = .51$, partial $\eta^2 = .008$.

The above analysis is based on using the filtering criterion of nonbelieved memories having memory ratings that were at least two points higher than belief ratings. When we adopted the criterion of memory ratings being at least three points higher than belief ratings, a statistically significant main effect of Belief was detected as well, $F(1, 55) = 44.10$, $p < .001$, partial $\eta^2 = .45$. As in Experiment 1, we performed an additional analysis by focusing on the items that participants recognized as "presented" in the recognition test.

Even when only the recognition responses with “Yes (presented)” were included, a statistically significant main effect of Belief was found, $F(1, 67) = 18.70$, $p <.001$, partial $\eta^2 = .21$, showing that undermining belief led to less efficient problem-solving behavior. No main effect of Memory Type was found, $F(1, 67) = 1.25$, $p =.26$. There was no significant main effect of Time Interval, $F(1, 67) = 0.003$, $p =.96$ and no statistically significant interactions were detected.[7]



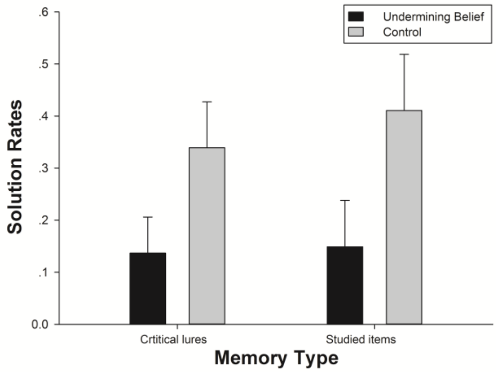


Figure 4.4a (left) and *4.4b* (right). Figure 4.4a illustrates solution rates primed by different Memory Types and Belief conditions in the immediate group (Experiment 2). Figure 4.4b illustrates results in the 1-week delay group. Error bars indicate 95% confidence intervals.

To rule out the possibility that the effect of belief was due to the difficulty of the CRATs, we changed the CRAT problems in the undermining belief and control conditions in around half (45.1%, $n = 32$) of the participants. We split the data into two groups in which participants received the opposite belief manipulation for the same materials. For instance, in one group, belief for “bread” was undermined and then participants solved a corresponding CRAT; in the other group, belief for “bread” was confirmed and participants solved the same CRAT. We conducted a 2 (Change: yes vs. no) × 2 (Memory Type: critical lures vs. studied items) ×2 (Belief: undermining vs. control) repeated measures ANOVA, with Change as a between-subject variable. No significant main effect of Change was found, $F(1, 55) = 2.44$, $p =.12$, indicating the materials did not impact our results.

[7] We examined whether being native English speaker or not would impact the results. When non-native English speakers were excluded, there were 29 participants in the immediate group and 17 participants in the delay group and the result pattern was not changed by the exclusion.

Solution times of CRATs

We were also interested in whether undermining belief would impact CRAT solution times. For critical items, we conducted a 2 (Time Interval: immediate vs. 1-week delay) ×2 (Belief: undermining vs. control) repeated measures ANOVA, with the first variable being between-subjects. Like Experiment 1, when participants solved no CRAT under the priming of (non)believed memories, no solution time data in that condition could be analyzed. There were 37 participants' solution time data that could be used. There was no main effect of Belief, $F(1, 36) = 0.45$, $p = .51$, Time Interval, $F(1, 36) = 0.28$, $p = .60$, or interaction effect, $F(1, 36) = 1.34$, $p = .25$. For studied items, we conducted the same analysis. We found no main effects of Time Interval or Belief. Thus, CRAT solution times for false and true memories were not impacted by Belief and Time interval.

Exploratory Analysis

We found no statistically significant correlation between compliance scores and number of nonbelieved memories (r (69) = .04, p = .73), which is consistent with Experiment 1. The data from both experiments suggest that compliance does not impact the formation of nonbelieved memories. Correlations between dissociative symptoms and number of true/false believed and nonbelieved memories were analyzed. We found no significant correlation between dissociative symptoms and true recognition rates (r (69) = -.21, p = 0.08), and also no significant correlation between dissociation and false recognition rate (r (69) = -.20, p = .09) emerged. No correlation was found between DES scores and overall number of nonbelieved memories (r (69) = .07, p = .59).

Consistent with Experiment 1, we found that undermining belief in false memories led to fewer CRAT problems being solved than the control condition. Belief is conceptualized as the truth value of an event. Even though a CRAT is an insight-based problem-solving task, retracting belief in false memories impacts the ability of false memories to prime CRATs. The results on false memories from Experiment 1 and Experiment 2 align with Bernstein et al.'s (2015) conclusion that false autobiographical beliefs, not memories, alter behavioral performance.

What we also found was that undermining belief in true memories resulted in lower CRAT solution rates. In Experiment 1, we did not find this. The reason might be that we made several improvements in Experiment 2, such as including more CRATs in each condition, and assigning the CRATs to each condition in a more balanced way. In addition, in Experiment 2 we found no statistically significant results on CRAT solution times. The main reason for this might be that there was limited data on the solution times for CRATs. Only when participants solved at least one CRAT problem could we obtain solution time data. This can be resolved in future studies by using easier CRAT problems. Also, our results showed that the deleterious effects of belief retraction on problem solving occurred both immediately and after one week. This shows that when belief is undermined, it does not have a short-lived effect, but it endures over time. This is line with research by Otgaar and colleagues (2013) who showed that nonbelieved memories can last for a month.

General Discussion

The current experiments serve as the first attempt to simultaneously assess the behavioral consequences of nonbelieved and believed memories on problem-solving behavior. We found evidence across two experiments that nonbelieved memories impacted problem-solving behavior under the conditions tested. The most intriguing finding was that undermining belief in false memories led to less efficient problem-solving behavior. This result persisted even after a 1-week delay. To our knowledge, this is the first experimental demonstration that for false memories, retracting belief adversely affects subsequent behavior.

This novel finding implies that belief contributed more to the behavioral performance on the CRATs than recollection. This is in line with previous research suggesting belief in the occurrence of past events was more important in determining eating behavior than recollection (Bernstein et al., 2015). However, in this work, no experimental test was performed to manipulate belief separately and examine whether this would impact behavior. Our study is the first showing the consequences of belief and recollection on behavior by using a non-inferential, and perhaps more automatic, insight-based problem-solving task.

Our findings have several theoretical implications. The data on nonbelieved memories support the distinction between belief and recollection. The distinction between belief and recollection was not made in the memory literature until recently (e.g., Otgaar et al., 2015; Scoboria et al., 2004; Scoboria et al., 2014; Mazzoni & Kirsch, 2002). Previous studies mostly showed dissociation of belief and recollection for autobiographical events. For instance, research has shown that belief in actions (e.g., clapping hands) could be undermined while the recollected aspect of the actions remained intact (Clark, Nash, Fincham, & Mazzoni, 2012; Mazzoni, Clark, & Nash, 2014). Also, experiments showed that belief in experiencing a hot balloon ride can be manipulated while recollection of the event remained intact (Otgaar et al., 2013). In our two experiments, we undermined belief for associatively-related words and recorded belief and memory ratings afterwards. We found that for both false and true memories, belief ratings dropped while memory/recollection ratings were high after undermining belief.

Scoboria et al. (2014) proposed a theoretical model to explain the relationship between belief and memory (recollection). In this model, autobiographical belief and recollection are two independent continuous dimensions that result in different categories such as believed memories and nonbelieved memories. Our experiments support this view inasmuch as we successfully manipulated participants' beliefs while recollections were retained. Based on the independence of these two components, it is proposed that belief in the occurrence of an event, rather than a specific memory for the event, is highly malleable and is the critical component in influencing behavior (Bernstein et al., 2015; Scoboria et al., 2014). Importantly, we found that for false memories, undermining belief led to a reduction in subsequent problem-solving behavior, a finding that accords well with the above proposition. As belief in occurrence is based in various inputs, just one of which is recollection, theories that focus on episodic recollection alone may not be the best predictors of behavior.

In Experiment 2, we found evidence that for true memory as well, problem-solving was more difficult when belief was undermined. This suggests that for memory in general, behavior is predominantly influenced by believing the event rather than recollecting the event. It is unclear why we did not find this effect in Experiment 1. Although one might expect that challenging true memories is more difficult than false memories (e.g., Otgaar, Candel, Smeets, & Merckelbach, 2010), this is not what we found in the present experiments. Here, both nonbelieved true and false memories were evoked, something that might be related to the fact that the DRM procedure leads both to high levels of true and false recognition, with false recognition rates often not differing from true recognition rates (Roediger & McDermott, 1995). Of course, future research should examine more closely whether belief is also important in guiding behavior for true memory.

Associative-activation theory (AAT, Howe, Wimmer, Gagnon, & Plumpton, 2009; Otgaar, Howe, Peters, Smeets, & Moritz, 2014), as well as the activation-monitoring theory (AMT, Roediger, Balota, & Watson, 2001) provide explanations for the priming effects of false memories. AAT suggests that processing of one concept activates a corresponding node and this activation spreads automatically to nearby associative concept nodes. When DRM list items are presented and encoded, their activation spreads to non-presented, but related items (i.e., the critical lure) resulting in false memories. Because false memories are highly associated to true memories, they often exert similar priming effect on CRAT problems (Howe et al., 2013). In our experiments, we manipulated participants' beliefs for the items after false and true memories were formed, and we found similar reduced priming effect for both true and false memories (Experiment 2). One possibility might be that undermining belief adversely affects spreading activation in one's knowledge base thereby reducing its effects on subsequent tasks including the spreading activation required to solve CRATs. Of course, further investigation is needed to examine the precise mechanism by which belief and recollection can impact problem-solving behavior.

One might argue that the manipulation of belief in our experiments might have changed the automatic nature of the priming process. In Experiment 2, the data showed that the CRAT solution rates primed by true memories still remained high even after a one-week interval, but previous research has found that these rates dropped after a one-week delay (Howe et al., 2013). Priming CRATs in prior research is considered to occur automatically by associative activation (Howe et al., 2010). Querying belief (e.g., undermining) in the current experiments may have made the recollections of true and false items more conscious and salient, which may have fundamentally changed the priming process. However, if we look at the data of the control condition in the immediate testing group, we found the exact same result with previous studies (solution rates: critical lures ⩾ studied items > unrelated items), and in the immediate group, belief was also queried in the control condition. This suggests that our belief manipulation might not have affected the automatic nature of our priming effects.

Our study also explored the relationship between compliance and nonbelieved memories. Both experiments demonstrated no statistical link between compliance scores and number of nonbelieved memories. Scoboria, Boucher, and Mazzoni (2015) found that the primary reason people retracted their belief in a memory was social feedback, such as someone telling you that your memory was not true. Our study found that people who were more compliant did not form more NBMs than people who were less compliant. One reason for this is that social feedback is more related to external pressure, such as suggestive information, whereas compliance can be regarded as an internal personality characteristic. Our null result begs the question whether the formation of NBMs might be more affected by external factors such as who provides social feedback (e.g., authority or stranger).

One might object that our memory task is related to the problem-solving task and that this is a potential confound in our experiments. However, in food preference studies (Bernstein & Loftus, 2009), participants created false beliefs or memories towards a negative food experience in the first session, and then the amount of that food they ate was measured in the second session. Participants' eating behavior was measured weeks or months after the first session (e.g., Geraerts et al., et al., 2008; Scoboria et al., 2012). They were told that the second session was a completely irrelevant experiment as to reduce the chance that participants could link the two sessions with each other. Importantly, in the current experiments, we also told participants that our problem-solving task was an unrelated task. By way of confirmation that our manipulation succeeded, we interviewed some participants after the experiments and none of them could see the link between the memory task and the problem-solving task.

To conclude, our experiments provide the first evidence that for false memories, problem-solving was hampered when belief was reduced. This shows that belief is the most active agent in impacting problem-solving behavior. Indeed, our experiments reveal novel evidence that belief and recollection have distinct behavioral consequences. The time has now come to extend this finding and investigate whether such differential consequences might also appear in other situations.

References

Bernstein, D. M., & Loftus, E. F. (2009). The consequences of false memories for food preferences and choices. *Perspectives on Psychological Science, 4,* 135–139.

Bernstein, E.M., & Putnam, E.W. (1986). Development, reliability, and validity of a dissociation scale. *Journal of Nervous and Mental Disease, 174*, 727–734.

Bernstein, D. M., Scoboria, A., & Arnold, R. (2015). The consequences of suggesting false childhood food events. *Acta Psychologica, 156,* 1–7.

Biondolillo, M. J., & Pillemer, D. B. (2015) Using memories to motivate future behavior: An experimental exercise intervention, *Memory, 23,* 390–402.

Blair, I. V., Lenton, A. P., & Hastie, R. (2002). The reliability of the DRM paradigm as a measure of individual differences in false memories. *Psychonomic Bulletin & Review, 9*, 590–596

Bowden, E.M., Jung-Beeman, M., Fleck, J., & Kounios, J. (2005). New approaches to demystifying insight. *Trends in Cognitive Sciences, 9,* 322–328.

Brewer, W. F. (1996). What is recollective memory? In D. C. Rubin (Ed.), *Remembering our past: Studies in autobiographical memory* (pp. 19–66). New York, NY: Cambridge University Press.

Clark, A., Nash, R. A., Fincham, G., & Mazzoni, G. (2012). Creating non-believed memories for recent autobiographical events. *PLOS One, 7,* 1–7.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58,* 17–22.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175–191.

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.

Giesbrecht, T., Lynn, S. J., Lilienfeld, S. O., & Merckelbach, H. (2008). Cognitive processes in dissociation: An analysis of core theoretical assumptions. *Psychological Bulletin*, *134*, 617–647.

Gudjonsson, G. H. (1989). Compliance in an interrogation situation: A new scale. *Personality and Individual Differences, 10*, 535–540.

Howe, M. L., Garner, S. R., Charlesworth, M., & Knott, L. (2011). A brighter side to memory illusions: False memories prime children's and adults' insight-based problem solving. *Journal of Experimental Child Psychology, 108*, 383–393.

Howe, M. L., Garner, S. R., Dewhurst, S. A., & Ball, L. J. (2010). Can false memories prime problem solutions? *Cognition, 117*, 176–181.

Howe, M. L., Garner, S.R., & Patel, M. (2013). Positive consequences of false memories. *Behavioral Sciences and the Law, 31*, 652–665.

Howe, M.L., Garner, S. R., Threadgold, E., & Ball, L.J. (2015). Priming analogical reasoning with false memories. *Memory & Cognition*, *43*, 879–895.

Howe, M.L., Threadgold, E., Norbury, J., Garner, S., & Ball, L.J. (2013). Priming children's and adults' analogical problem solutions with true and false memories.. *Journal of Experimental Child Psychology*, *116*, 96–103.

Howe, M. L., Wilkinson, S., Garner, S. R., & Ball, L. J. (2016). On the adaptive function of children's and adults' false memories. *Memory, 24*, 1062–1077.

Howe, M. L., Wilkinson, S., Monaghan, P., Ball, L. J., & Garner, S. R. (2013). Why are illusions adaptive? The curious case of false memories. *Manuscript submitted for publication.*

Howe, M. L., Wimmer, M. C., Gagnon, N. & Plumpton, S. (2009). An associative activation theory of children's and adults' memory illusions. *Journal of Memory and Language, 60*, 229–251.

James, W. (1950). *The principles of psychology* (Vols. 1 and 2). Mineola, NY: Dover Publications. (Original work published 1890)

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.

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.

Mazzoni, G.A.L., Clark, A., & Nash, R.A. (2014). Disowned recollections: Denying true experiences undermines belief in occurrence but not judgments of remembering. *Acta Psychologica, 145*, 139–146.

Mazzoni, G. A. L., & Kirsch, I. (2002). False autobiographical memories and beliefs: a preliminary metacognitive model. In T. Perfect, & B. Schwartz (Eds.), *Applied* *metacognition* (pp.121–145). Cambridge, UK: Cambridge University Press.

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

Otgaar, H., Candel, I., Smeets, T., & Merckelbach, H. (2010). “You didn’t take Lucy’s skirt off”: The effect of misleading information on omissions and commissions in children’s memory reports. *Legal and Criminological Psychology, 15*, 229–241.

Otgaar, H., Howe, M. L., Peters, M., Smeets, T., & Moritz, S. (2014). The production of spontaneous false memories across childhood. *Journal of Experimental Child Psychology, 121*, 28–41.

Otgaar, H., Howe, M. L., van Beers, J., van Hoof, R., Bronzwaer, N., & Smeets, T. (2015). The positive ramifications of false memories using a perceptual closure task. *Journal of Applied Research in Memory and Cognition, 4*, 43–50.

Otgaar, H., Moldoveanu, G., Wang, J., & Howe, M.L. (2016). *Exploring the consequences of nonbelieved memories in the DRM paradigm.(under review).*

Otgaar, H., Scoboria, A., & Mazzoni, G. (2014). On the existence and implications of nonbelieved memories. *Current Directions in Psychological Science, 23,* 349–354.

Otgaar, H., Scoboria, A., & Smeets, T. (2013). Experimentally evoking nonbelieved memories for childhood events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 717–730.

Pezdek, K., & Salim, R. (2011). Physiological, psychological and behavioral consequences of activating autobiographical memories. *Journal of Experimental Social Psychology, 47,* 1214–1218.

Roediger, H., Balota, D., & Watson, J. (2001). Spreading activation and arousal of false memories. In H. Roediger, J. Nairne, & A. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder. Science conference series* (pp. 95–115). Washington, DC: American Psychological Association.

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,* 803–14.

Rubin, D. C. (2006). The basic-systems model of episodic memory. *Perspectives on Psychological Science, 1*, 277–311.

Schacter, D. L. (1996). *Searching for memory: The brain, the mind and the past*. New York, NY: Basic Books.

Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. London, UK: Chapman & Hall.

Scoboria, A., Boucher, C., & Mazzoni, G. (2015). Reasons for withdrawing belief in vivid autobiographical memories. *Memory, 23,* 545–562.

Scoboria, A., Jackson, D., Talarico, J., Hanczakowski, M., Wysman, L., & Mazzoni, G. (2014). The role of belief in occurrence within autobiographical memory. *Journal of Experimental Psychology: General, 143,* 1242–1258.

Scoboria, A., Mazzoni , G. , Jarry, J. , & Bernstein, D. (2012). Personalized and not general suggestion produces false autobiographical memories and suggestion-consistent behavior . *Acta Psychologica*, *139*, 225–232 .

Scoboria, A., Mazzoni, G., Kirsch, I. & Relyea, M. (2004). Plausibility and belief in autobiographical memory. *Applied Cognitive Psychology, 18,* 791–807.

Scoboria, A., & Talarico, J. M. (2013). Indirect cueing elicits distinct types of autobiographical event representations. *Consciousness and Cognition, 22*, 1495–1509.

Scoboria, A., Talarico, J. M., & Pascal, L. (2015). Metamemory appraisals in autobiographical event recall. *Cognition*, *136,* 337–349.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology*, *26*, 1–12.

Appendix 4: Examples of DRM lists and CRAT problems

Critical Lures	Black	Bread	Car	Needle	Fruit	Shirt
	white	butter	truck	thread	apple	Blouse
	dark	Food	bus	pin	vegetable	sleeves
	charred	Eat	automobile	eye	orange	pants
	night	sandwich	vehicle	sewing	kiwi	tie
	funeral	Rye	drive	sharp	citrus	button
	colour	Jam	jeep	point	ripe	shorts
	grief	Milk	Ford	prick	pear	iron
	death	flour	keys	thimble	banana	polo
	ink	Jelly	garage	haystack	berry	collar
	coal	dough	highway	thorn	cherry	vest
	brown	crust	van	injection	basket	pocket
	grey	Slice	taxi	syringe	juice	jersey
Associated CRAT	Board/mail /magic	Crust/stale /french	Chase/ police/toy	Knitting/ pine/work	Salad/ bowl/juice	Football/ flannel/vest

CHAPTER 5

Undermining Memory Changes Decision Making

This Chapter is an adapted version of the following paper:

Wang, J., Otgaar, H., Smeets, T, Howe, M. L., & Zhou, C. (under revision). Manipulating memory associations changes decision-making preferences in a preconditioning task.

Abstract

Memories of past experiences can guide our decisions. Thus if memories are undermined or distorted, decision making should be affected. Nevertheless, little empirical research has been done to examine the role of memory in reinforcement decision-making. We hypothesized that, if memories guide choices in a conditioning decision-making task, manipulating these memories would result in a change of decision preferences to gain reward. We manipulated participants' memories by providing false feedback that their memory associations were wrong before they made decisions that could lead them to win money. Participants' memory ratings decreased significantly after receiving false feedback. More importantly, we found that false feedback led participants' decision bias to disappear after their memory associations were undermined. Our results suggest that reinforcement decision-making can be altered by manipulating memories. The results are discussed using memory mechanisms such as spreading activation theories.

Keywords: memory, false feedback, decision-making, reinforcement learning, sensory preconditioning

In the early 1930s, one of Pavlov's dogs demonstrated the sensory preconditioning effect (see Kimmel, 1977). A whistle and a light were paired together several times, after which the dog was conditioned to flex its limb (using electric shock) upon presentation of the light. This resulted in the whistle also eliciting limb flexion, even though the whistle was never reinforced (see also Brogden, 1939). The dog obviously formed the "whistle-light" association as well as the "light-shock" association, and somehow integrated these two memory associations to guide its reaction towards the whistle. This reaction implies that memory mechanisms might play a role in reinforcement learning/decision-making. After all, if the dog could not remember the whistle-light association, it probably would not flex its limb to the whistle.

Remembering previous experiences enables organisms to make decisions to acquire reward or avoid harm. For example, people would pick a restaurant with a good memory in which they have enjoyed the food. Or people with memories of getting allergic with peanuts would avoid eating peanuts again. Surprisingly, little empirical research has been done regarding how memory impacts reinforcement decision-making, and only recently the role of episodic memory in reinforcement decision-making has attracted scientific attention (see Gershman & Daw, 2017 for a recent review; see also Weber & Johnson, 2006). Classical reinforcement learning theories usually adopt the computational perspective such as a statistical summary of past experiences (Shohamy & Daw, 2015). For instance, by making a bet on number 8, a person sometimes wins money (e.g., 80 times out of 100 times) and sometimes loses money (e.g., 20 times out of 100 times). The computational approach would assume that the person will bet on number 8 again because the summarised winning probability of betting on number 8 is 80%. This approach, however, has ignored how the experiences might be encoded in people's mind. For instance, as in selective forgetting (Anderson & Hanslmayr, 2014; Lind, Visentini, Mäntylä, & Missier, 2017), the person might selectively remember only the times that he has won, which could inflate his confidence in winning; or he merely remember the times that he has lost, which would prevent him from betting more.

Indeed, Ludvig, Madan, and Spetch (2015) have recently found that priming memories of past wins can induce risk-seeking. They showed that participants who received a reminder of a winning experience tended to choose a risky choice more often than those who received no reminders. More recently, Bornstein, Khaw, Shohamy, and Daw (2017) were the first to illustrate how consulting individual memories can bias reinforced decision-making. In their task, participants chose between two slot machines to win money and learned by trials and errors that each slot machine delivered winning or losing tickets at different probabilities. At a later stage, participants were presented with the tickets as memory reminders before they made choices to win money. The results showed that those reminders of past choices strongly influenced participants' choices afterwards in that a reminder of a past action with reward made participants repeat the action while a reminder of loss made them avoid the action.

Neurobiological research has suggested the importance of memory associations in reinforced decision-making as well. Wimmer and Shohamy (2012) have firstly shown that reward decision-making was correlated with activations in the hippocampus, where memory associations are usually formed. In their task similar to Pavlovian's sensory preconditioning task, participants first learned that a picture (S1+) and a circle always appeared together, and in a reward phase later they learned that the circle led to monetary reward. As a result, participants displayed a preference to choose the S1+ picture to win money above other equally familiar stimuli, even though it had never been rewarded. Neural imaging data revealed that this decision preference was predicted by activity in the hippocampus, a brain area encoding memory associations. This research showed that reward can spread across memory associations (e.g., from the rewarded circle to associated S1+ picture) via memory circuits in the hippocampus, which then impacts decision making. Our research question is, if memories for associations can guide reward decision-making, is it possible to alter decision making by changing participants' memory associations?

Memory Malleability and False Feedback

Episodic memory refers to mentally re-experiencing one's own previous experiences and it binds different elements of "what", "where" and "when" into an integrative experience (Tulving, 2002). More than 80 years of false memory research has demonstrated that memory is a highly adaptive and re-constructive system where its elements can be selectively sampled, intentionally forgotten or even completely distorted (Howe & Otgaar, 2013; Loftus, 2005; Schacter, 2012). For instance, people watched a parade of eight people, but then misremembered there were six people after answering a suggestive question "was the leader of four people a male?" (Loftus, 1975). Or one can falsely recall that a blued-eye person committed a crime after receiving suggestive false information, while the perpetrator actually had brown eyes (e.g., Zajac & Henderson, 2009).

A plethora of research has demonstrated that memory can be manipulated via providing false feedback. One type of manipulation is to create rich autobiographical false memories through false feedback when in fact there are no relevant experiences. After being falsely suggested that they had been sick after eating a particular food such as egg salad in childhood, some participants report remembering vivid details of being sick on egg salad as a child and later show reduced preference and eating of that food (Bernstein, Laney, Morris, & Loftus, 2005; Geraerts, Berstein, Merckelbach, Linders, Raymaekers, & Loftus, 2008; Scoboria, Mazzoni, & Jarry, 2008; Wang, Otgaar, Bisback, Smeets, & Howe, 2017). In those studies, participants have never experienced the egg salad-sickness association, yet they exhibit preference changes when they come to believe and remember the association.

A more recent research line has used false feedback to undermine or "weaken" genuine episodic memories. In Mazzoni, Clark, and Nash's (2014) study, participants performed actions such as clapping their hands and rubbing the table in front of a video camera. A few days later, their memories of the performed actions were tested and false feedback was provided to some of the actions that those actions were never performed. Some participants no longer believed that they had performed these actions but still remembered them, referring to the formation of nonbelieved memories. Importantly, some memory characteristics, such as spatial and temporal clarity, even became weaker after false feedback. Other studies have found that false feedback led people to no longer believe in their memories of studied words and later changed memory related priming effects and problem-solving performance (Otgaar, Moldoveanu, Wang, & Howe, 2017; Wang, Otgaar, Howe, Smeets, Merckelbach, & Nahouli, 2017). Here, participants have experienced performing actions/seeing words, but false feedback to their memories make them disbelieve in their memories and change their memory-related behavior.

As we have summarized above, false feedback to experiences can either create a completely false memory or make people retract a memory. As we have reviewed in the beginning, recent research has shown that memories about previous experiences play an important role in reinforcement decision making such as gambling. Thus it is possible to manipulate people's memories for past experiences and then change people's choices generated from conditioning or reinforcement. Recent studies have revealed that people's choices can be changed by reminders of relevant memories (i.e., selectively sampled memory) (Ludvig et al., 2015; Bornstein et al., 2017). However, to our knowledge, no study has used false feedback for true memories to alter reinforcement decision making. The current study aims to study whether false feedback to memory associations would change participants' choices to gain reward in a Pavlovian preconditioning task. We hypothesized that changing memories of past experiences (i.e., how experiences are encoded in people's mind) would impact reinforcement decision making.

The Present Study

Combining findings from memory research with the reward decision-making paradigm (e.g., Wimmer and Shohamy, 2012), in the current study, we used false feedback to change participants' memory associations that particular items were paired together, and then examined whether such false feedback would affect participants' choices among items to win money. To be more specific, participants first learned associations among pictures and circles (e.g., an S1+ picture always associated with a yellow circle) and then learned that certain circles (e.g., the yellow circle) led to winning of money. In a memory test later, participants were told that the S1+ picture did not appear together with the yellow circle, but was associated with another non-rewarded circle. Finally, they were asked to choose from pictures and circles to gain money.

A recent model of memory suggests that autobiographical memory can be dissociated into two components, namely *recollection* referring to mental re-experiencing of events with sensory details and *belief* that refers to the truth value subjectively attached to a memory (Rubin, 2006; Scoboria, Jackson, Talarico, Hanczakowski, Wysman, & Mazzoni, 2014). For instance, people believe their birth but do not have recollections of the birth event or someone remembers vivid images of Santa Claus but does not believe in it anymore. Plenty of studies have demonstrated that belief in a memory can be changed or undermined via feedback, in which for example participants are told that their memory is incorrect (e.g., Mazzoni, Clark, & Nash, 2014; Otgaar, Wang, Franken, & Howe, 2017). Such feedback can lead people to disbelieve that particular experiences happened though their sense of recollection is sustained (see Otgaar, Scoboria, & Mazzoni, 2014 for a review). After false feedback was provided in our study, recollection and belief of memory associations were measured in the memory test, and we predicted that participants' memory associations would be undermined in terms of these memory aspects, especially their beliefs in the associations (Mazzoni et al., 2014; Otgaar et al., 2014).

Theories of reinforcement learning and decision making have recently attempted to incorporate episodic memories into animal and human learning frameworks. For instance, Weber and Johnson (2006) have proposed in their preference-as-memory (PAM) theory that the structure of memory representations are very relevant for decision making and different structures of memory may be the root of some phenomena in decision making. One novelty with our procedure is that we did not change memory for the *stimuli*, but manipulated the memory *association* between two stimuli. In memory theories such as spreading activation theories (Anderson, 1983; Howe, Wimmer, Gagnon, & Plumpton, 2009; Roediger, Balota, & Watson, 2001), memory associations between memory nodes are key structures in an associative memory network, which determines the direction of spreading memory activation. If participants' self-reported memory associations do contribute to reinforcement decision making, by undermining memory associations between the S1+ and rewarded circle, we would expect participants' decision preferences for S1+ pictures to disappear after false feedback. On the other hand, if reinforcement decision making is driven by mechanisms other than memory, we would expect participants' decision-making preferences unchanged by memory manipulations.

Method

Participants

Before recruiting participants, we used G*Power 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007) to calculate the required sample size. With an estimated power of .80, power analysis revealed that 41 participants were required to detect a medium effect size (d = 0.45). Forty-one students from Fudan University, Shanghai participated in our study for a financial reward of ¥30. The sample consisted of 7 males and 34 females, with age ranging from 18 to 27 years old (M_{age} = 21.2, SD = 2.22).

Design and Procedure

The current study was a within-subject design. In the False feedback condition, participants received false feedback on their memorized associations in order to break the established associations. In the True feedback (control) condition, participants received true feedback on their memory associations. Half of the associations were provided with false feedback and the other half with true feedback. We also included filler pictures (no associations formed) in the task, which served as a baseline of decision preferences. The procedure basically followed the same steps as in the sensory preconditioning paradigm by Wimmer and Shohamy (2012), except that we included a memory test with feedback before the decision phase.

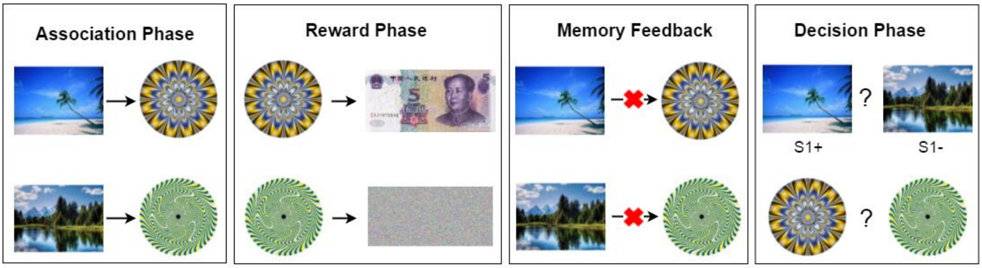


Figure 5.1. Demonstration of procedure in the False feedback condition. The True feedback condition had an identical procedure except that the memory feedback given was true. Figure 5.1 illustrates one category of pictures (scenes). In total there were four categories of pictures (see Appendix).

Association Phase. As Figure 5.1 shows, in the association phase, neutral pictures were paired with neutral patterned circles. Participants were merely asked to view some pictures in the instruction. However, without telling them to memorize associations, a picture (scene, furniture, body part or vehicle) always appeared before a particular patterned circle (see Walther, 2002). There were four categories of pictures and in each category there were two pictures: one picture (labeled as S1+ picture) paired with a later rewarded circle and the other picture (labeled as S1- picture) paired with a non-rewarded circle. Order of the pairs was randomized and each pair was presented for 10 times randomly. Four filler picture pairs (picture-picture) appeared randomly as well. Each picture or circle was presented for 1.5 s, with 1 s interval between them. The interval between each pair was 3.5 s to help participants learn the associations. After all pairs were presented, participants rated their likings for all stimuli on 1-7 scales (1 = very disliked, 7 = very liked).

Reward Phase. Following the association phase, participants learned in the reward phase that half of the circles led to monetary reward while the other half led to no reward. They were instructed to respond with the “F” key within 500 ms when they saw a picture of a money bill and respond with the “J” key when detecting a grey square. Once they responded in time and correctly, they won the money to their virtual money account. The money bill was always preceded by half of the circles while the grey square was preceded by the other half. Thus, there was a 100% contingency rate in the reward phase. Participants were explicitly told to use the circle cues to respond so that they could learn the reward associations.

Feedback Phase. Before the decision phase, participants completed an incidental memory test for the associations in the association phase. The memory test first asked participants to recognize which circle was paired with a particular picture (two choices were provided: a correct one and a wrong one). Then feedback was provided immediately after each recognition. For half of the associations (n = 4; two S1+ associations and two S1- associations), the correct answers were purposively exchanged for S1+ and S1- stimuli in the E-prime program. So for those associations, the program falsely told them that the correct answer of an S1+ stimulus was a non-rewarded circle, while it told participants that the correct answer of an S1- stimulus was a rewarded circle. At the same time, the experimenter told the participant: “Your memory is wrong. I saw clearly that the picture was paired with this circle (pointing to the other answer).” For the other half of the associations (n = 4), the program provided the true status of associations while giving feedback. Four filler picture pairs were also included in the memory test, and they were always provided with correct feedback to make the overall feedback credible. After feedback was given, participants rated their recollection (“Do you actually remember that the two items were paired together?”; 1 = no memory at all, 8 = complete memory) and belief (“Do you believe that the two items were paired together?”; 1 = do not believe at all, 8 = definitely believe) on 1-8 scales for the original association (see Scoboria et al., 2014).

Decision Phase. Finally, participants went through the decision phase. For each trial, two pictures or two circles appeared left and right on the screen. Participants were asked to choose the picture that they thought would win them money. Each trial consisted of two pictures from the same category (e.g., beach vs. lake or leg vs. arm), and one picture (S1+ picture) had been associated with a rewarded circle while the other (S1- picture) had been associated with a non-rewarded circle in the association phase. The rewarded circle and non-rewarded circle were presented in another trial to assess reward learning. In the decision phase, the same two stimuli were presented for four times, with each stimulus randomly assigned to the left or right side. There were 4 categories × 2 pairs × 4 times = 32 critical trials in total, which were presented in random order.

Two filler pictures that were never paired with any circles were also presented together with S1- pictures in the decision phase, to assess a baseline non-associated decision preference. Note that as decision choice might be impacted by the stimuli, materials were counterbalanced in a way that S1+ and S1- stimuli were exchanged in around half of the participants (53.6%; n=22). Pre-liking data also revealed no significant liking difference for S1+ (M= 4.34, 95%CI [4.05, 4.63]) and S1- (M= 4.23; 95%CI [3.94, 4.51]) stimuli, t (41) = 1.01, p = .32.

Results

Baseline memory data and manipulation check

In the memory test, participants first were asked to choose the patterned circle that they recalled was associated with the picture. Memory accuracy for associations pre-false feedback (M= 0.68, 95%CI [0.57, 0.79]) did not differ significantly from the memory accuracy for associations pre-true feedback (M= 0.77, 95%CI [0.65, 0.89]), t (40) = -1.13, p=.27, indicating equivalent levels of associative memories in the two conditions pre-feedback.

After feedback provided in the memory test, participants rated their recollections and beliefs for the S1+ associations. As Figure 5.2 shows, false feedback led to statistically lower recollection (t (40) = -5.26, p < .001, Cohen's d = - 0.84) and belief ratings (t (40) = -7.11, p< .001, Cohen's d = - 1.11) than true feedback. False feedback manipulation worked at both belief and recollection dimensions, suggesting that it substantially weakened memory associations. Using a cut point of 4.5 to dichotomously categorize recollection versus no recollection and belief versus no belief in 1-8 scales, Figure 5.2 shows that mean ratings for associations in the true feedback condition fell into the upper range (4.5-8) while mean ratings for associations in the false feedback condition fell into the lower range (1-4.5).

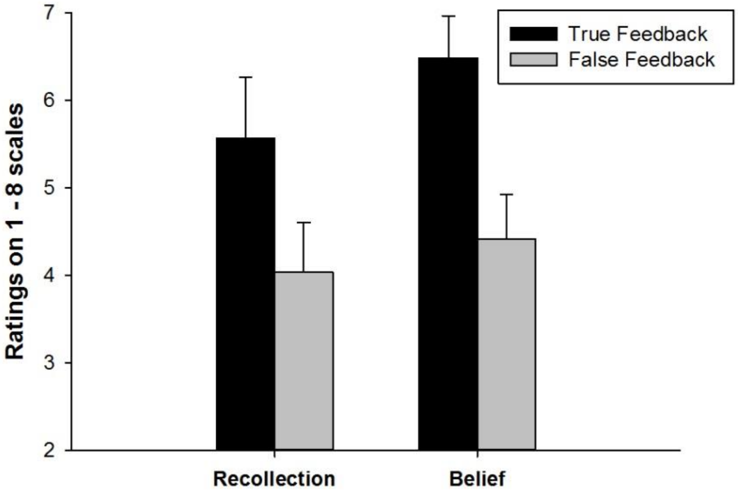


Figure 5.2. Recollection and Belief ratings after True and False feedback. Error bars represent 95%CI.

Decision-making rates

Since preconditioning effect is limited by the conditioning effect (i.e., reward learning), data per comparison (S1+ vs. S1-) in the decision phase were included only when participants learned the corresponding reward association. For example, if a participant did not learn which circle would win them money in the first place, there was unlikely any spreading of preference from the circle to the picture. Five participants did not learn any reward association in the False feedback condition (not greater than chance level) and four participants did not learn any reward association in the True feedback condition. Thus, thirty-two participants' data were included in the final set of analyses. To be noted, the results did not differ with or without exclusion of participants. [8] Post-hoc power analysis calculated the power as 0.83 when sample size was 32. Those participants chose on average 97.66% (95%CI [0.96, 1.00]) of the rewarded circles in the False feedback condition and

[8] Even if we included all 41 participants data without controlling for reward learning, similar results were achieved: False feedback condition had significantly lower rate of choosing S1+ than True feedback condition, t (40) = -2.24, p = .03, Cohen's d = - 0.55. However, controlling reward learning is cleaner, more honest way of analyzing the data (i.e., the effect might be due to different levels of reward learning), and thus we mainly report the results based on the conditionalized data.

96.48% (95%CI [0.94, 1.00]) of the rewarded circles in the True feedback condition. Paired samples *t*-test revealed no statistically significant difference, t (31) = 0.83, p = .41, implying similar (near ceiling) reward learning in the two conditions.

We were most interested in the decision making on S1 stimuli when true and false feedback were provided to participants' memory associations. We calculated the average rate of choosing S1+ over S1- in four rounds for each participant as per Wimmer and Shohamy's (2012). In the baseline decision-making condition (filler vs. S1-), participants chose the filler pictures at an approximately chance level (M = 0.49, 95%CI [0.36, 0.62]). In the True feedback condition, participants chose S1+ in 66.41% (95%CI [0.54, 0.79]) of the choices, which was statistically larger than 50% chance level, $t(31)= 2.62$, p = .01, and statistically larger than the baseline condition, $t(31)= 2.80$, p = .009, Cohen's d = 0.50. Thus we have replicated the sensory preconditioning effect in the true feedback condition. However, in the False feedback condition, participants chose S1+ only in 40.23% (95%CI [0.28, 0.53]) of the choices, which was statistically lower than the True feedback condition, $t(31)= - 2.74$, p = .01, Cohen's d = - 0.74, and not different from the baseline condition, $t(31)= - 0.98$, p = .33.

Decision-making preference scores

To better illustrate decision preferences over S1+ or S1- in the True vs. False feedback conditions, we calculated participants' average decision preference scores. Decision preference score for pictures was calculated by subtracting times of choosing S1- stimuli from times of choosing S1+ stimuli over four rounds. A positive value indicates that participants preferred choosing S1+ stimuli over S1- stimuli; a negative value indicates participants preferred S1- stimuli; 0 value means no preference over S1+ or S1- stimuli (at 50% chance level). The preference score ranged from - 4 to 4. For example, a value of 4 means that the participant chose 4 times of S1+ and 0 time of S1- picture. Paired samples *t*-test showed that the mean preference score in the True feedback condition (M= 1.44, 95%CI [0.47, 2.40]) was significantly higher than that of the False feedback condition (M= -0.78, 95%CI [-1.80, 0.24]), t (31) = -2.99, p = .005, Cohen's d = - 0.53. Figure 5.3a demonstrates that True feedback led participants to preferably choose S1+ stimuli (i.e., the preconditioning effect), while false feedback led them to show no preference over S1+ or S1- stimuli (i.e., the preconditioning effect disappeared).

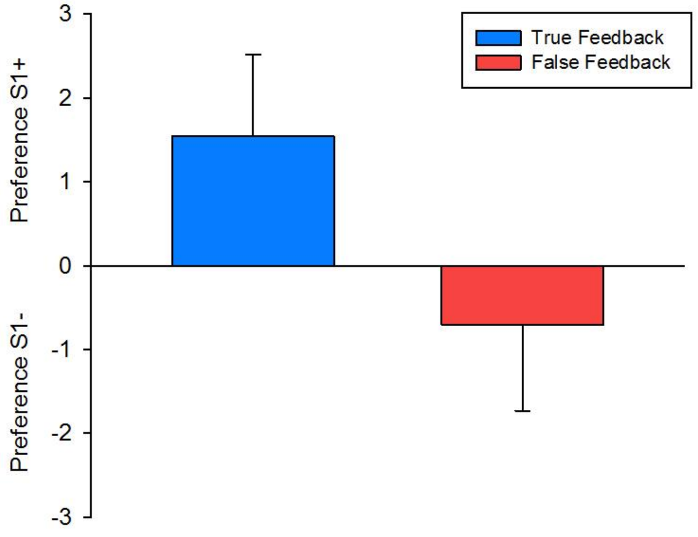


Figure 5.3a. Mean decision preferences in False and True feedback conditions. Decision preference score = (Times of choosing S1+) – (Times of choosing S1-). E.g., a value of 0 means no preference and a value of 2 means participants averagely chose 2 more times of S1+ in four rounds (i.e., choosing S1+ 3 times and choosing S1- once). Error bars represent 95%CI.

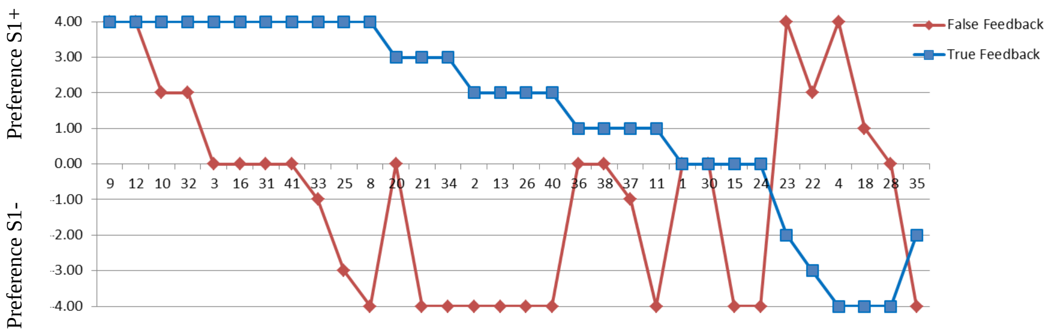


Figure 5.3b. Individual data of decision preferences in False and True feedback conditions. Each number on the X-axis represents one participant.

A recent study suggests that individual data should be presented with small sample sizes (Weissgerber, Milic, Winham, & Garovic, 2015). Figure 5.3b illustrates individual data of decision preference in true and false feedback conditions. It shows that decision preference was even reversed in the False feedback condition compared to the True feedback condition at an individual level. When a participant preferred S1+ in the True feedback condition, this participant preferred S1- (or not preferring S1+) in the False feedback condition; and when a participant preferred S1- in True feedback condition, she/he was more likely to choose S1+ stimuli in the False feedback condition. We separated participants into two groups, those who preferred S1+ stimuli in the True feedback condition (i.e., preference score > 0; n =22) and those who preferred S1- stimuli in True feedback condition (i.e., preference score < 0; n = 6). Note that all these participants learned the reward conditioning very well (*Mean rate of choosing rewarded circles* = 97%; 95%CI [94%, 100%]). Paired samples t-test compared decision preferences between False versus True feedback in each group: for the former group, t (21) = -7.90, p < .001, Cohen's d = - 1.96; for the latter group, t (5) = 3.20, p = .02, Cohen's d = 1.39.

Decision-making latencies

Reaction times (RTs) for S1+ versus S1- choices were recorded during the decision phase. After RTs < 300ms or > 10000 ms were excluded (2.02% of RTs data) based on general rules in reaction time tasks such as the implicit memory test (Greenwald, Nosek, & Banaji, 2003), RTs were averaged in each condition for each participant. A paired samples t-test was conducted to examine whether false feedback impacted decision-making speed. No statistically significant difference in reaction times was found between False feedback (M = 1802 ms, 95%CI [1539, 2066]) and True feedback (M = 1782 ms, 95%CI [1494, 2071]) conditions, t (40) = 0.17, p = .86.

Discussion

This is the first study to manipulate memory associations in a reinforcement decision-making task. The results showed that false feedback to participants' memories resulted in decreased recollection and belief ratings for their memory associations. That is, after being told that they misremembered two items paired together, participants indicated that they recollected the association to a lesser extent and less believed in their memory associations than when receiving true feedback. The false feedback manipulation thus successfully weakened or broke participants' memory associations, suggesting memory associations between stimuli can be reconstructed.

Previous research on false feedback and memory has mostly focused on how feedback can change people's beliefs in a memory while recollections are well retained (Otgaar, Scoboria, & Mazzoni, 2014). For example, developmental psychologist Jean Piaget remembered vividly that he was kidnapped as a child but later he was told by his nurse that she fabricated the story. He no longer believed in his memories but he could recall vivid recollections of the kidnap event (Mazzoni, Scoboria, & Harvey, 2011). In our study, participants not only reported lower beliefs in their memory associations but also reported lower recollections of the memory associations after receiving false feedback. The reason might be that memory associations formed in the current study were temporal in nature. That is, the presence of an S1 picture always preceded the presence of a circle to form an association. Mazzoni et al. (2014) found that, unlike sensory details, time characteristics of a memory can be significantly impacted with false feedback, which explains why temporal associations in our study can be easily reconstructed.

This is also the first study showing that manipulating how memory associations are represented in people's minds can impact reinforcement decision-making. Intriguingly, the most novel finding is that false feedback changed participants' decision preferences compared to true feedback. Participants exhibited the classical sensory preconditioning effect after receiving true feedback in that they chose S1+ stimuli more often over S1- stimuli to gain money. However, after they received false feedback telling them that certain associations were unfounded, they no longer preferred S1+ stimuli and showed no sensory preconditioning effect anymore. Individual data showed that some participants even preferred S1- to gain monetary reward in the false feedback condition. Likely because we switched the paired circles for S1+ and S1- stimuli in the false feedback, some participants might have formed a false memory that the S1- picture was associated with a rewarded circle and thus demonstrated preferences over S1- stimuli instead of showing no preference during decision making.

Recent research has shown that reminders of particular memories can greatly guide people's decision making (Ludvig et al., 2015; Bornstein et al., 2017). Our study showed that undermining memories about past experiences can change reward-based decision making. All these studies suggest that how organisms remember their experiences (at the time of making a decision) is a key mechanism underlying decision making or general reward learning. More specifically, our results can be readily explained by memory mechanisms such as spreading activation theories (Anderson, 1983; Howe et al., 2009; Roediger et al., 2001). From a memory network perspective, memory consists of mental representations of stimuli (i.e., nodes) that can be associated with each other. In the true feedback condition that exhibited a typical sensory preconditioning effect, an S1+ stimulus was encoded to associate with a circle that was encoded to associate with a reward. Thus, associations were established between an S1+ picture and a circle and also between the circle and reward in a participant's memory network. According to the spreading activation principle, activation of a node can spread along its associations to nearby nodes automatically (Anderson, 1983; Howe et al., 2009). Although the S1+ stimulus was never rewarded, the activation of the reward can be automatically spread via the rewarded circle to the S1+ stimulus during decision making.

However, by providing false feedback to participants' memories, their memories were re-assessed and probably re-constructed as numerous research has shown the flexibility and malleability of our memory system (e.g., Loftus, 2005; Schacter, 2012). Indeed, our results showed that participants' memories for the associations between S1+ stimuli and rewarded circles were significantly undermined. Thus when an S1+ picture was presented in the decision phase, the value of reward might not be activated since the association between S1+ and the rewarded circle was weakened. As a result, participants did not show any preference to S1+ stimuli. Previous research has also generally discussed the role of memory structure and spreading activation in decision making (e.g., Morewedge & Kahneman, 2010; Weber & Johnson, 2006). Our explanation shares a lot of similarity with Morewedge and Kahneman's (2010) proposition that principles of associative memory activation should apply to judgment processes. They propose that strongly activated information is likely to gain more weight than information that is not activated in associative memory network during decision making, which can explain why reminders of a specific memory (i.e., stronger activation) determine later decision making. In the current experiment, we have successfully manipulated participants' memory associations to change their decision preferences, which aligns with the principle of spreading activation along memory associations.

Taken together, the current study has implications for theories of reinforcement decision-making in general. Research on reinforcement learning usually focuses on algorithms such as the model-free and model-based approaches, where organisms compute possible values for different choices and choose the option leading to greatest expected value (O'Doherty, Cockburn, & Pauli, 2017). Gershman and Daw (2017) pointed out that the computational perspective is limited in mimicking situations relevant for real life (e.g., facing a situation that has never been encountered), and they have proposed an episodic reinforcement learning model where episodic memory serves as a nonparametric estimation of the value in the algorithms. The PAM theory also suggests that decisions are made by retrieving information from memory in order to determine the best option (Weber & Johnson, 2006). Although the current study examined the role of memory from a different approach by manipulating memories, the results support the recent theoretical view that episodic memory should be integrated within the reinforcement decision-making system. That is, principles of memory (e.g., memory activation, false memory effect, etc.) may account for decision making. For example, participants who were implanted with a false memory that they used to get sick with peach yogurt showed lower preferences for peach yogurt (Scoboria et al., 2008). Further research is needed to investigate the role of memory in decision making.

Some might argue that an inclusion of a memory test would change the nature of the sensory preconditioning task. However, results in the true feedback condition still revealed a sensory preconditioning effect while the baseline condition revealed no effect, which suggests that the inclusion of a memory test did not impact the sensory preconditioning effect. Some might wonder whether the experimental effect was due to demand characteristics that participants purposively selected S1+ stimuli in the true feedback condition and purposively showed no preference in the false feedback and baseline conditions. There are several reasons why this is unlikely to be the case. There were a lot of different trials (112 trials in total) mixed from different conditions in the decision phase, and we have also switched the S1+ and S1- stimuli in half of the participants. During the decision-making phase, both the experimenter and the participant did not know exactly which stimuli were S1+ or S1- and which stimuli had received true or false feedback (one can only know by checking the program). What's more, Wimmer and Shohamy's (2012) study suggests that the transference of value from a rewarded circle to the S1+ picture is out of participants' awareness. Participants in our study made a decision relatively fast (in around 1.8 s) and RTs in the true feedback condition did not differ from RTs in the false feedback condition. Nevertheless, it is unknown from the current study whether participants had explicitly referred to their memories or if the spreading of preference via associations was automatic. Further research is needed to closely examine participants' decision-making process in the preconditioning task.

To conclude, our research has presented novel evidence that false feedback undermined participants' memory associations in a preconditioning task and false feedback to their memories led to participants no longer showing decision preferences to gain reward. The results suggest that episodic memories play an important role in reinforcement decision making. It is time now to investigate how principles of memories may impact reinforcement decision making.

References

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior, 22*, 261-295.

Anderson, M. C., & Hanslmayr, S. (2014). Neural mechanisms of motivated forgetting. *Trends in Cognitive Sciences, 18*, 279-292.

Bernstein, D. M., Laney, C., Morris, E. K., & Loftus, E. F. (2005). False beliefs about fattening foods can have healthy consequences. *Proceedings of the National Academy of Sciences of the United States of America, 102*, 13724-13731.

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.

Bernstein, D. M., Scoboria, A., & Arnold, R. (2015). The consequences of suggesting false childhood food events. *Acta Psychologica, 156*, 1–7.

Bornstein, A. M., Khaw, M. W., Shohamy, D., & Daw, N. D. (2017). Reminders of past choices bias decisions for reward in humans. *Nature Communications, 8*, 15958.

Brogden, W. J. (1939). Sensory pre-conditioning. *Journal of Experimental Psychology, 25*, 323-332.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191.

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.

Gershman, S. J., & Daw, N. D. (2017). Reinforcement learning and episodic memory in humans and animals: An integrative framework. *Annual Review of Psychology, 68*, 101-128.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85*, 197-216.

Howe, M. L., & Otgaar, H. (2013). Proximate mechanisms and the development of adaptive memory. *Current Directions in Psychological Science, 22*, 16-22.

Howe, M. L., Wimmer, M. C., Gagnon, N. & Plumpton, S. (2009). An associative activation theory of children's and adults' memory illusions. *Journal of Memory and Language*, *60*, 229-251.

Kimmel, H. D. (1977). Notes from" Pavlov's Wednesdays": sensory preconditioning. *American Journal of Psychology*, *90*, 319-321.

Lind. M., Visentini, M., Mäntylä, T., & Del Missier, F. (2017). Choice-supportive misremembering: A new taxonomy and review. *Frontiers in Psychology, 8*, 2062.

Loftus, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology, 7*, 560-572.

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

Ludvig, E. A., Madan, C. R., & Spetch, M. L. (2015). Priming memories of past wins induces risk seeking. *Journal of Experimental Psychology: General*, *144*, 24.

Mazzoni, G.A.L., Clark, A., & Nash, R.A. (2014). Disowned recollections: Denying true experiences undermines belief in occurrence but not judgments of remembering. *Acta Psychologica, 145*, 139–146.

Morewedge, C. K., & Kahneman, D. (2010). Associative processes in intuitive judgment. *Trends in Cognitive Sciences, 14*, 435-440.

O'Doherty, J. P., Cockburn, J., & Pauli, W. M. (2017). Learning, reward, and decision making. *Annual Review of Psychology*, *68*, 73-100.

Otgaar, H., Moldoveanu, G., Wang, J., & Howe, M. L. (2017). Exploring the consequences of nonbelieved memories in the DRM paradigm. *Memory, 25*, 922-933.

Otgaar, H., Scoboria, A., & Mazzoni, G. (2014). On the existence and implications of nonbelieved memories. *Current Directions in Psychological Science, 23,* 349–354.

Otgaar, H., Scoboria, A., & Smeets, T. (2013). Experimentally evoking nonbelieved memories for childhood events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 717–730.

Otgaar, H., Wang, J., Fränken, F., & Howe, M. L. (2017). Not Believing in Autobiographical Experiences: Social Feedback and External Evidence Affects Belief in Occurrence, but not Belief in Accuracy and Recollection. Manuscript submitted for pulication.

Roediger, H. L., Balota, D., & Watson, J. (2001). Spreading activation and arousal of false memories. In H. Roediger, J. Nairne, & A. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder. Science conference series* (pp. 95–115). Washington, DC: American Psychological Association.

Rubin, D. C. (2006). The basic-systems model of episodic memory. *Perspectives on Psychological Science, 1*, 277-311.

Schacter, D. L. (2012). Adaptive constructive processes and the future of memory. *American Psychologist, 67*, 603-613.

Scoboria, A., Jackson, D., Talarico, J., Hanczakowski, M., Wysman, L., & Mazzoni, G. (2014). The role of belief in occurrence within autobiographical memory. *Journal of Experimental Psychology: General, 143*, 1242-1258.

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

Shohamy, D., & Daw, N. D. (2015). Integrating memories to guide decisions. *Current Opinion in Behavioral Sciences, 5,* 85-90.

Tulving, E. (2002). Episodic memory: from mind to brain. *Annual Review of Psychology, 53*, 1-25.

Walther, E. (2002). Guilty by mere association: evaluative conditioning and the spreading attitude effect. *Journal of Personality and Social Psychology, 82*, 919-934.

Wang, J., Otgaar, H., Bisback, A., Smeets, T., & Howe, M. L. (2017). The consequences of false recollections and beliefs on food preferences. Annual conference of European Association of Psychology and Law, Mechelen, Belgium, 2017.

Wang, J., Otgaar, H., Howe, M. L., Smeets, T., Merckelbach, H., & Nahouli, Z. (2017). Undermining belief in false memories leads to less efficient problem-solving behavior. *Memory, 25,* 910-921.

Weber, E. U., & Johnson, E. J. (2006). Constructing preferences from memory. In Lichtenstein, S. & Slovic, P. (Eds.), *The Construction of Preference (pp. 397–410).* Cambridge University Press.

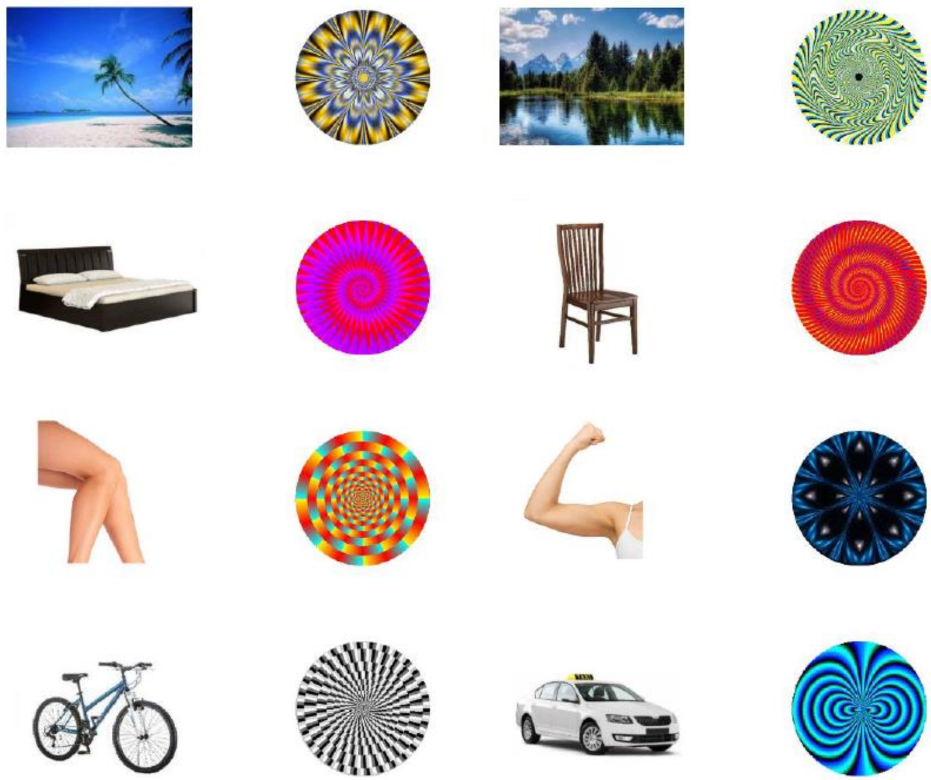
Weissgerber, T. L., Milic, N. M., Winham, S. J., & Garovic, V. D. (2015). Beyond bar and line graphs: time for a new data presentation paradigm. *PLoS biology*, *13*, e1002128.

Wimmer, G. E., & Shohamy, D. (2012). Preference by association: how memory mechanisms in the hippocampus bias decisions. *Science, 338*, 270-273.

Zajac, R., & Henderson, N. (2009). Don't it make my brown eyes blue: Co-witness misinformation about a target's appearance can impair target-absent line-up performance. *Memory*, *17*, 266-278.

Appendix 5: Materials

S1+	Rewarded circles	S1-	Non-rewarded circles



CHAPTER 6

Belief, Not Recollection, Impacts Food Preferences

This Chapter is an adapted version of the following paper:

Wang, J., Otgaar, H., Bisback, A., Howe, M. L., & Smeets, T. (under review). The consequences of implicit and explicit beliefs on food preferences.

Abstract

Memories can have consequences on, for instance, people's eating behavior. In the current experiments, we first examined the effect of belief versus recollection on people's food preferences, and then investigated whether explicit belief (i.e., self-reported) or implicit belief (i.e., measured by an autobiographical implicit association test; aIAT) had a similar effect on food preferences. Participants (N = 163) were falsely told that they got sick after eating egg salad in their childhood and then received guided imagery to induce false recollections/beliefs concerning the food-aversive event. Half of the participants with false memories were debriefed and told that the event was false in order to reduce their belief in the event. We found that belief, not recollection regarding the food-aversive event impacted participants' food preferences (Experiment 1a). Furthermore, we found that explicit, but not implicit, belief predicted participants' food preferences (Experiment 1b). The current results suggest that explicit judgments of belief in a memory may explain the behavioral consequences resulting from memories.

Keywords: Autobiographical belief, recollection, implicit belief, food preference

False memories can exert significant consequences on people's attitudes and behavior. For instance, false memories of being sexually abused in childhood could lead to a lawsuit against a family member (Loftus, 1993). When people recall those false memories with vivid sensory details, they oftentimes also believe that the recalled event really happened and thus act consequently upon their memory. However, the question arises regarding which aspect of memory is more likely to contribute to behavioral consequences: the vivid recollections of the experience or a mere belief in the occurrence of the event? For example, when people do not believe in the occurrence of an event but still hold a vivid recollection, do they still behave in a manner consistent with this nonbelieved memory? In the current experiments, we were interested in the effect of belief on behavior and whether this effect is similar at an implicit and explicit level.

Belief versus Recollection

The relevance of this question becomes evident when we consider the fact that autobiographical recollection and belief are two independent aspects that contribute to the experience of autobiographical memory (Mazzoni, Loftus, & Kirsch, 2001; Scoboria, Jackson, Talarico, Hanczakowski, Wysman, & Mazzoni, 2014; Scoboria, Talarico, & Pascal, 2015). *Recollection* is the feeling that we are re-experiencing the event accompanied by mental images of a memory that can include sensory details and references to vividness and context. *Belief* is regarded as the truth value of the event; that is, whether it happened or not (Rubin, 2006; Scoboria et al., 2014). We have recollections and beliefs for most of our memories, but for some events, we only have autobiographical beliefs with no recollections as, for example, when we believe that we were born but cannot recall it; or we have only recollections but no beliefs which have been referred to as nonbelieved memories (NBMs; Mazzoni, Scoboria, & Harvey, 2010, Otgaar, Scoboria, & Mazzoni, 2014).

Over the past decade, many studies have demonstrated that false beliefs for a particular food can change food preferences and eating behavior (e.g., Bernstein, Laney, Morris, & Loftus, 2005; Bernstein & Loftus, 2009; Geraerts, Bernstein, Merckelbach, Linders, Raymaekers, & Loftus, 2008; Scoboria, Mazzoni, Jarry, & Bernstein, 2012). In a prototypical study of this kind, participants receive false feedback that they had been ill after eating a particular food (e.g., peach yogurt, egg salad, asparagus), and then, participants who falsely believe the suggested event exhibit reduced preference for that food and eat less of that particular food even weeks or months later (Geraerts et al., 2008; Scoboria, Mazzoni, & Jarry, 2008).

One problem with these previous studies is that it sometimes was unclear whether recollection or belief drove the food preferences effect. To give an example, studies sometimes used the term "false belief" to refer to both notions of false memory and false belief (e.g., Berkowitz, Laney, Morris, Garry, & Loftus, 2008; Bernstein, Laney, Morris, & Loftus, 2005; Mantonakis, Wudarzewski, Bernstein, Clifasefi, & Loftus, 2013), but sometimes used the term "false memory" referring to both false memory and false belief (see Bernstein & Loftus, 2009; Laney, Morris, Bernstein, Wakefield, & Loftus, 2008). As a demonstration, Bernstein et al. (2005) suggested to participants that they had been ill after eating strawberry ice cream in childhood and found decreased preference for strawberry ice cream in participants categorized by them as believers, those who either falsely remembered or falsely believed the event had occurred. Although the majority of believers (82%, n = 41) simply believed but did not remember the false event, it is unclear to what extent recollection contributed to the food preference effect observed in Bernstein et al.'s study.

More recently, a few studies have attempted to study the unique contribution of false memories and false beliefs to food preferences. Scoboria, Mazzoni, Jarry, and Bernstein (2012) found that participants who reported a false memory about the suggested food-aversive event for peach yogurt ate statistically less peach yogurt afterward than both the false beliefs only group and the control group and the latter two groups did not differ. However, Bernstein, Scoboria, and Arnold (2015) conducted a mega-analysis on eight studies published from 2005 to 2008, and found that those who remembered the suggested event were indistinguishable from people who had only beliefs in terms of food preferences, suggesting false beliefs are more important in impacting people's attitudes and behavior. One important limitation from previous studies, however, is that they did not experimentally dissociate belief from recollection in order to better test the separate effects of belief and recollection on food preferences.

Research has shown that debriefing participants by telling them that they had false memories, participants will relinquish their beliefs in those memories resulting in nonbelieved memories (e.g., Clark, Nash, Fincham, & Mazzoni, 2012; Otgaar, Scoboria, &

Smeets, 2013; Wang, Otgaar, Howe, Smeets, Merckelbach, & Nahouli, 2017). Social feedback challenging people's memories is the most influential factor in manipulating or relinquishing beliefs in memories (Scoboria, Boucher, & Mazzoni, 2015). To dissociate the effects of recollection and belief on food preferences, recollection (yes vs. no) and belief (yes vs. no) need to be orthogonally differentiated to examine their consequences. Hence, four groups should be formed respectively: a false memory group (recollection and belief), a false belief group (belief but no recollection), a NBM group (recollection but no belief), and a control group (no recollection and no belief). The aim of the current experiment was to address the methodological limitation in previous studies by creating these four groups and examining their effects on food preferences. The results of Experiment 1a only address the question on how explicit belief affects food preferences.

Explicit versus Implicit Belief

Besides measuring people's explicit belief with self-reported questionnaires, researchers have also successfully measured people's implicit belief (or implicit truth value) of an event using the autobiographical Implicit Association Test (aIAT; Agosta & Sartori, 2013; Sartori, Agosta, Zogmaister, Ferrara, & Castiello, 2008; Shidlovski, Schul, & Mayo, 2014). The aIAT measures participants' reaction times towards contrasting autobiographical events. For instance, in one study, guilty participants went through a mock crime where they stole a CD containing an upcoming exam while innocent participants did not experience it. Later, guilty and innocent participants reacted to statements such as "I stole the CD" and "I did not steal the CD". Guilty participants responded faster when true statements were associated with "I stole the CD" than when true statements were associated with "I did not steal the CD" (Sartori et al., 2008). Agosta and Sartori (2013) reviewed recent studies and found that the aIAT has a 90% accuracy rate when detecting implicit truthfulness of a memory.

It has been found that implicit measurements by the IAT can affect behavior. McConnell and Leibold (2001) measured white undergraduates' implicit attitudes towards black people and tested their interactions with black experimenters. Their results showed that those who had negative implicit attitudes toward blacks had more negative interactions with a black (vs a white) experimenter. Other researchers found that implicit preferences over political parties were good predictors of future voting behavior, even though people claimed no explicit preferences over the parties (Arcuri, Castelli, Galdi, Zogmaister, & Amadori, 2008).

However, there is a debate regarding the predictive validity of the IAT measurements recently. Greenwald, Poehlman, Uhlmann, and Banaji (2009) conducted a meta-analysis on the predictive validity of the IAT on actual behavior and found an average r = .27 for the correlations between IAT and behavioral and physiological measures. They also found an average r = .36 for the prediction of behavior by explicit self-reported measures. However, Oswald, Mitchell, Blanton, Jaccard, and Tetlock (2013) examined the reviewed studies more closely and they found that the average correlation between IAT and behavior was r = .15, much lower than Greenwald et al. (2009) found. Oswald and colleagues found that IATs were poor predictors of every criterion of discriminatory behavior and IATs were no better predictors than explicit measures. So far, the predictive validity of IATs is still under discussion and no consensus has been reached (see Greenwald, Banaji, & Nosek, 2014 and Oswald, Mitchell, Blanton, & Jaccard, 2015 for more discussion).

Considering the above-mentioned studies, it is unclear whether implicit belief measured by aIAT would predict people's behavioral indices, such as food preferences. For example, if a participant implicitly holds a belief that she had been sick after eating egg salad, will she exhibit low preference for egg salad? We suggested to participants that they had been sick after eating egg salad in their childhood, which has been shown to effectively induce explicit beliefs and memories in that autobiographical event (Bernstein et al., 2015). We were also interested in using the autobiographical IAT to measure the implicit belief of the suggested event in different belief and recollection groups. Previous studies on IAT usually focused on the relationship between people's *implicit attitudes* and behavior (Greenwald et al., 2009). To our knowledge, no study has used aIAT to detect the *implicit beliefs* in suggested events and no study has examined whether implicit belief can impact behavior directly. Our goal was to use the aIAT (Sartori et al., 2008) to measure participants' implicit belief in the suggested event of being sick on egg salad and examine whether implicit belief impacted food preferences over egg salad.

The Current Experiments

We used an adapted *false feedback paradigm* (Bernstein et al., 2005; Scoboria et al., 2008) to induce false recollections and beliefs for a suggested food event that the subject was sick after eating egg salad in their childhood. Participants first completed a battery of questionnaires online to measure their baseline data such as autobiographical memories, food history, and food preferences. Two weeks later, they came to the lab and were falsely told that they had been sick after eating egg salad in their childhood. Then participants went through a guided imagery session to retrieve memories of the suggested event, something that is known to induce high levels of rich false memories (Garry & Wade, 2005; Scoboria et al., 2012).

After the guided imagery, participants were asked to indicate whether they had recollections of the suggested event. A novel procedure of the current study was that, for half of the participants who recalled false memories, the experimenter debriefed them with the truth (i.e., that the suggested event was fabricated and that the profile was fake) to undermine their beliefs in their memories and thus induce nonbelieved memories (recollection but no belief). For other participants who were not debriefed, they naturally fell into the following categories after the false suggestion: those who had recollections and beliefs, those who had beliefs but no recollections, and those who had no recollections and no beliefs. Finally, all participants performed the aIAT to measure their implicit belief for the suggested event and were also asked about their food preferences and eating intention.

Experiment 1a examined how autobiographical belief or recollection might impact people's food preferences. Bernstein, Scoboria, and Arnold's (2015) mega-analysis revealed that belief is more important than recollection in impacting people's behavior such as food preferences. Recently, Wang et al. (2017) also found that belief in memories is crucial when changing problem-solving efficiency. We predicted that, if autobiographical belief is more important than recollection, we would find an effect of belief but no effect of recollection in impacting participants' food preferences. Experiment 1b examined whether implicit belief measured via an aIAT would impact people's food preferences relative to explicit belief as well as the relationship among implicit belief, recollection and explicit belief.

Experiment 1a

Method

Participants

Participants were screened via a Food History Inventory (FHI) that measured people's experiences with different foods. To be included in this study participants were required to not have a history of being sick on egg-salad as measured by the FHI. One hundred and sixty-three participants passed the screening and were tested in our study. Two participants were subsequently excluded because it was indicated that they definitely remember or believe the sick-on-egg salad event in the online Autobiographical Memory and Belief Questionnaire. The final sample consisted of 36 males and 125 females ranging from 18 to 39 years old (M=22.24, SD= 3.49). Participants received course credits or 7.5 euros shopping vouchers. The study was approved by the ethical review committee of the Faculty of Psychology and Neuroscience, Maastricht University.

Materials

Food History Inventory (FHI; Bernstein et al., 2005). The 24-item food history questionnaire asks participants about their experience with food, such as *ate too much ice cream* and *helped someone peel potatoes*. Participants indicated whether these events happened to them or not in the past on a 1-8 scale (1= definitely did not happen, 8= definitely did happen). The inventory included the critical item *felt sick after eating egg salad*.

Autobiographical Belief and Memory Questionnaire (ABMQ; Scoboria et al., 2004). The ABMQ assesses participants' autobiographical belief and memory for five past events. The five events include *lost while shopping, sick on egg salad, enjoyed egg salad, broke a window with a hand* and *sick on peach yogurt*. The belief and recollection items for the event *sick on egg salad* were of main interest for the current study. Participants rated each item for each event on a 1-8 point Likert-like scale. For the belief item, participants were asked "How likely is it that you personally, before the age of 7, were sick after eating egg salad?" (1= definitely did not happen, 8= definitely did happen). For the recollection item, participants were asked "Do you actually remember being sick on eating egg salad before you were the age of 7?" (1= no memory of event at all, 8 = clear and complete memory).

Memory Characteristic Questionnaire (MCQ; Johnson, Foley, Suengas, & Raye, 1988). The questionnaire measures phenomenal characteristics associated with memory on 1 to 7 scales, including visual details, smell, taste, occurred setting, time and negative emotions associated with the critical event.

Memory/Belief Form (MvB). The MvB was adapted from the Memory/Belief form in Bernstein, Laney, Morris, and Loftus (2005). We asked participants to indicate whether they had a memory (i.e., recollection; yes or no) and whether they had a belief (yes or no) that they had got sick after eating egg salad in childhood (see also Otgaar, Scoboria, Howe, Moldoveanu, & Smeets, 2016).

Food Preferences Inventory (FPI; Bernstein et al., 2005). Participants rated 62 foods on 1–7 point scales (1= definitely don't like; 7= definitely like). We included the target food – egg salad, and several related items (other varieties of salads, yogurts, etc.) to assess participants' preferences for foods.

Restaurant Questionnaire. This measurement of behavioral intention was adapted from the restaurant questionnaire used by Laney et al. (2008). Participants were told that a restaurant targeting student consumers in Maastricht was conducting some simple marketing research. Participants rated 25 foods on a Likert-like scale from 1 (definitely no) to 8 (definitely yes) to indicate willingness to order various foods at that restaurant, including the critical item *toasted bread with egg salad*.

Procedure

Session 1

Participants completed an online questionnaire via Qualtrics that took about 15 to 20 minutes. In the questionnaire, they were informed that people tend to forget what happened to them between the age of 3 and 7 years old. The cover story was that researchers were examining whether it was possible to recover memories before the age of 7. Participants were told that their data collected by the online questionnaire would be analyzed to identify events that very likely occurred in their childhood. Participants completed their demographic data, the FHI, the ABMQ and the FPI. The FHI was used to screen participants who had no history of being sick after eating egg salad in childhood. The

ABMQ measured participants' baseline data about memory and belief ratings on the critical event (sick on egg salad in childhood). The FPI measured participants' baseline data about their preference concerning different foods including the critical food egg salad. Eligible participants were invited to take part in session 2 two weeks later.

Session 2

Session 2 occurred two weeks after Session 1 and took about 30 minutes. Participants were invited to the lab to complete the experiment. Similar with previous studies (Bernstein & Loftus, 2009; Scoboria et al., 2008), participants were (falsely) told that a professional health company helped analyze the data from their online questionnaire and generated a profile of *Food Statistic Assessment*. The experimenter took out the profile from an envelope and participants were asked to read the profile. The cover page of this (fake) profile indicated the name, birth date, nationality and test date of the participant. The profile included the participant's preference concerning different kinds of foods from the online questionnaire in Session 1 measured by the FHI and FPI. The key information in the profile was a page that listed possible personal childhood events and *being sick after eating egg salad* was listed with highest probability (7 out of 8 on a 1-8 scale). We asked participants to think about clues that would support the critical event for several seconds and we would try to recover their memories regarding this event.

Following this, participants received a guided imagery session in an attempt to elaborately recall the event. The script for guided imagery was similar to that of Scoboria et al. (2012). During the guided imagery, participants were first asked to close their eyes, relax, and mentally go back to the time when they were little by imagining for instance how they looked like, with whom they played and where they went to school. Next, they were asked to focus on the details of the critical event. Specifically, they were guided to focus on the environment, the food, the smell, the taste and finally the feelings they might experience of being sick on egg salad. Participants were suggestively told that several symptoms such as cramps in the stomach and nauseous might have occurred during this event, and they were free to report any feelings or anything that popped up to their imagery. The guided imagery session was audiotaped. The duration of guided imagery including preparation was around 8 minutes.

After the guided imagery, participants completed the MvB and the MCQ. Participants were first asked whether they recalled a memory (recollection) by circling "yes" or "no" and rated the memory strength on a 1-8 scale. We explained that by "memory" we meant they could recall the time, the place and the content of the event and that they thus could mentally re-live the event. Their memory statements were recorded by the experimenter. Also, the MCQ was used to verify the phenomenology of their memories. Importantly, since our purpose was to dissociate belief from recollection and to examine their consequences, a random half of the participants who reported memories were debriefed by the experimenter before measuring their beliefs by telling them: "*I have to debrief you that the true purpose of our study is to induce false memories. The event that you were sick after eating egg-salad actually never happened—this is indicated in the answers of your online survey*." Other participants did not receive the debriefing but were just asked to confirm whether they had memories or not. Then all participants filled in the belief questions: whether they had a belief and their belief strength on a 1-8 scale.

Upon completion of the MvB and MCQ, all participants went through the aIAT[9] for about 10 minutes. Finally, they completed the Food Preferences Inventory and the Restaurant Questionnaire. The Restaurant Questionnaire was pretended as an extra marketing research.

Results

Recollection and Belief

Participants were classified as having recollection versus no recollection and belief versus no belief primarily based on self-reported responses in the MvB. To be noted, participants were told that they needed to recall specific details and report those details to the experimenter if they had any specific recollections. Forty-one percent of the participants (n = 66) reported having vivid recollections and the majority 59% (n = 95) reported no memory. The false recollections reported were relatively rich in details. For instance, one participant reported that "*I was at a mother's friend's house. There was a big bowl at the table; there was a pink dressing on it, a lot of eggs. In the beginning it was disgusting, too moisty. So I put something else in my mouth*". Another typical report was, "*Sunday evening all at home. My mom made egg salad. I ate too much and too fast. I feel stomach ache. I complained to my mom the next day that I cannot go to the kindergarten*".

[9] The procedure and results of aIAT are reported in Experiment 1b.

The experimenter recorded participants' recollections, and half of those with recollections (n=33) were debriefed by the experimenter that their memories were false. After debriefing, 17 participants indicated "no" belief and another 4 participants with memories indicated "no" belief spontaneously (without debriefing). Thus, a total of 21 participants had NBMs. Table 6.1 shows the number of participants who circled "yes" or "no" on the recollection and belief questions of MvB in the end. For participants with no recollections after receiving the false suggestion, 26 participants believed the critical event happened and 69 indicated no belief in the occurrence of the critical event.

Table 6.1
Number of participants with/without recollection or belief in the critical event after debriefing

		Recollection		
		Yes	No	Total
Belief	Yes	45 (28%)	26 (16%)	71 (44%)
	No	21 (13%)	69 (43%)	90 (56%)
	Total	66 (41%)	95 (59%)	161

To verify participants' self-reported recollections, we examined their ratings on the MCQ of the critical event. All participants were asked to rate the phenomenological characteristics when they thought about the critical event. As Figure 6.1 shows, participants who reported recollections had the highest phenomenological scores on all memory characteristics.

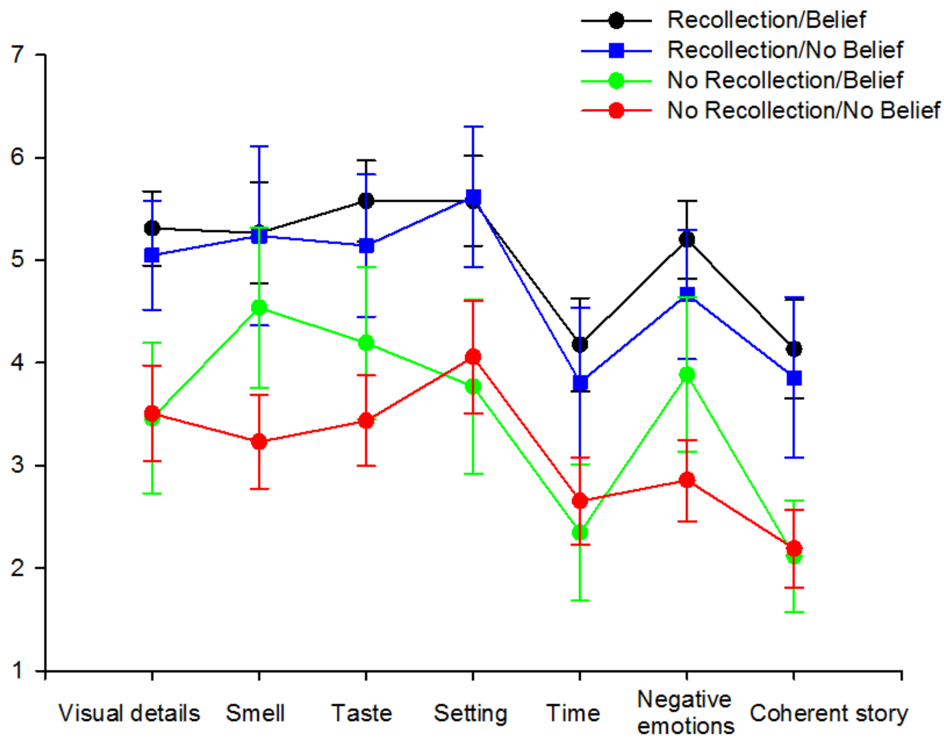


Figure 6.1. Phenomenological characteristics of the critical event in different Recollection (yes vs. no) and Belief (yes vs. no) combination groups. Error bars stand for 95% confidence intervals.

We also examined the recollection and belief strength (1-8 scales) of the critical event pre- and post- suggestion. Note that in the online questionnaire of Session 1, belief was measured twice, and we averaged them as the baseline belief scores. Table 6.2 reveals the mean recollection and belief ratings with 95% CIs pre- and post-suggestion in each group. As we can see from the table, for the Recollection/Belief group, recollection and belief ratings were both high (above 6) in Session 2 and the mean belief rating did not differ from the mean recollection rating, $t(44) = 0.71$, $p = .48$. For the Recollection/No Belief group, recollection ratings were significantly higher than belief ratings after the manipulation in Session 2, $t(20) = 2.43$, $p = .03$, $d = 0.53$. The mean difference between recollection and belief ratings was 1.62, which also satisfied the criterion of NBMs based on rating differences (Scoboria & Talarico, 2013). For the No Recollection/Belief group, belief ratings were significantly higher than memory ratings, $t(25) = 6.67$, $p < .001$, $d = 1.33$.

Table 6.2
Mean Recollection and Belief ratings in Session 1 (pre-suggestion) and Session 2 (post-suggestion) in different groups (CI: confidence interval)

		No Recollection/ No Belief (*n*=69)	No Recollection/ Belief (*n*=26)	Recollection/ No Belief (*n*=21)	Recollection/ Belief (*n*=45)
Recollection Rating (*M*, 95%CI)	Session 1	1.17 [1.02, 1.33]	1.50 [0.89, 2.11]	1.71 [1.07, 2.36]	1.98 [1.42, 2.53]
	Session 2	1.94 [1.68, 2.21]	3.12 [2.44, 3.79]	5.29 [4.52, 6.05]	6.02 [5.60, 6.45]
Belief Rating (*M*, 95%CI)	Session 1	2.08 [1.84, 2.32]	3.01 [2.21, 3.82]	2.55 [1.92, 3.18]	2.61 [2.24, 2.98]
	Session 2	2.72 [2.43, 3.02]	5.35 [4.82, 5.87]	3.67 [2.86, 4.47]	6.16 [5.76, 6.56]

Food Preferences

We were most interested in examining whether recollection or belief impacted food preferences. Since participants were categorized in 2 (Recollection: yes vs. no) × 2 (Belief: yes vs. no) by the MvB in the end, we conducted a 2 (Recollection) × 2 (Belief) between-subjects ANCOVA to analyze preferences on egg salad, controlling for baseline preference scores over egg salad in Session 1. Levene's test indicated equal variance across groups, $F(3, 157) = 0.01$, $p > .99$, and the criteria for ANCOVA were also met. There was no significant interaction effect between Recollection and Belief, $F(1, 156) = 1.10$, $p = .30$. We found a statistically significant main effect of Belief[10], $F(1, 156) = 6.76$, $p = .01$, partial $\eta^2 = .04$, Cohen's $d = 0.58$, suggesting that belief in being sick on egg salad lowered food preference over egg salad (see Figure 6.2a). However, no main effect of Recollection was found, $F(1, 156) = 0.55$, $p = .46$, implying recollection of the critical food event did not impact preference over egg salad.

[10] Although we had unequal sample sizes in different conditions, the variance is equal across groups. To eliminate possible confounding from unequal sample sizes, we ran a regression model with dummy variables from the independent variables. We found the same results as in ANCOVA that only Belief explains participants' preference scores over egg salad.

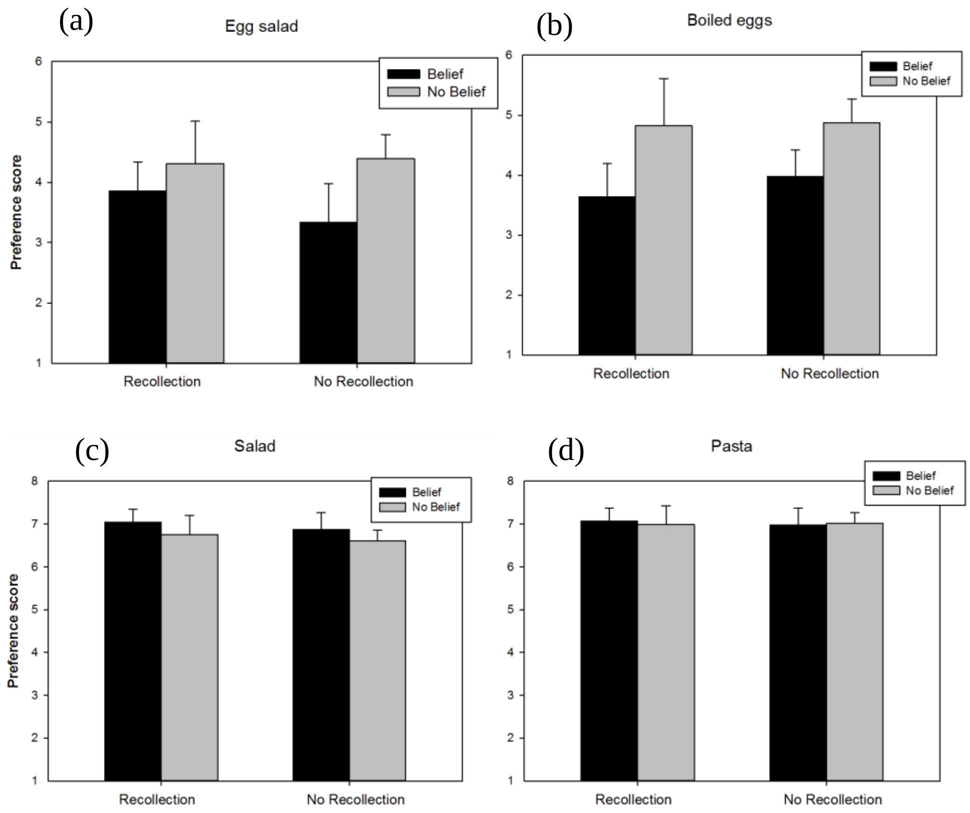


Figure 6.2. Food preferences over (a) Egg salad, (b) Boiled eggs, (c) Salad, and (d) Pasta in different Recollection (yes vs. no) and Belief (yes vs. no) groups. (Error bars 95%CIs)

Participants' food preference ratings on other foods were measured as well (e.g., boiled eggs, salad, watermelon, etc.). We conducted 2(Recollection) × 2(Belief) between-subjects ANCOVAs to test participants' preferences over other foods (controlling for baseline preference scores). Interestingly, participants' preferences over boiled eggs were impacted in a similar way as preference for egg salads. We found a significant main effect of Belief, $F(1, 142) = 10.32$, $p = .002$, partial $\eta^2 = .07$, Cohen's $d = 0.65$, but no significant main effect of Recollection, $F(1, 142) = 0.37$, $p = .54$, and no interaction effect, $F(1, 142) = 0.20$, $p = .66$. Furthermore, neither Recollection nor Belief in the egg salad event impact participants' preferences over other irrelevant foods such as salad in general, watermelon, and pasta (see Figure 6.2c and 6.2d).

Eating Intention

Participants' behavioral intention on eating egg-salad in a restaurant was measured by asking them how likely they would like to order toasted bread with egg-salad in a restaurant on a 1-8 scale. Since the restaurant questionnaire was only administered in Session 2, no baseline data in Session 1 was recorded. Levene's test showed equal variance across groups, $F(3, 157) = 0.23$, $p = .88$. We conducted a (Recollection: yes vs. no) × 2 (Belief: yes vs. no) ANOVA on eating intention of egg salad, but did not detect any significant main effect of Belief, $F(1, 157) = 2.09$, $p = .15$, or main effect of Recollection, $F(1, 157) = 0.24$, $p = .62$, and no interaction effect, $F(1, 157) = 2.21$, $p = .14$. We conducted exploratory group comparisons with Bonferroni correction thereby showing a statistical difference between the Belief/No Recollection group ($M = 3.08$, 95%CI [2.26, 3.89]) and the No Belief/No Recollection group ($M = 4.16$, 95%CI [3.66, 4.66]), $p = .03$, Cohen's $d = 0.51$.

Discussion

We induced false recollections and beliefs for a childhood event that one had been sick after eating egg salad. A non-trivial proportion (41%, $n = 66$) of participants produced false recollections and 44% ($n = 71$) of all the participants ended up with believing the suggested event. Most importantly, we found out that belief in the critical event, not recollection of the event, impacted participants' food preference scores on egg salad and boiled eggs. When participants believed that they used to be sick after eating egg salad, they exhibited lower preferences in eating egg salad and boiled eggs compared to those who did not have a belief in the egg salad event. Recollections, on the other hand, no matter how rich they were, did not impact participants' preferences in egg salad and boiled eggs.

The findings are consistent with Bernstein et al.'s (2015) conclusion that belief might be more important in impacting people's attitudes and behaviors. We extended their finding in that, when people did not believe in the critical food event, even with recollections, people's food preferences were not impacted either. Belief is the truth value attached to a memory that indicates whether the recollections represent what truly happened. When people assign a "false" tag (i.e, no belief) to a memory, it indicates that their recollections are not an honest representation of what happened, thus there is no need to act upon the false recollections anymore.

We did not find an effect of belief on eating intention. We only found that when there was no recollection, belief in the critical event led to lower intention to order egg salad in a restaurant compared to no belief. One possible reason might be because we did not control for eating intention in Session 1 but merely measured participants' eating intention in Session 2. Without controlling the baseline eating intention, the no-difference among groups might be due to the individual differences among groups. Another possible reason is that eating intention is different from food preferences, thus it might be more difficult to be affected by belief.

Having established that belief, and not recollection impacts food preferences, we were interested in investigating whether it is explicit (i.e., self-reported) belief or implicit belief that impacted food preferences. We have reported those results in Experiment 1b where we measured participants' implicit belief (using aIAT) in the critical event that one had been sick after eating egg salad after participants received false suggestion. The aIAT has been shown to be highly accurate in detecting implicit truthfulness for autobiographical experiences (Agosta & Sartori, 2013). We first examined if participants' implicit belief in the critical event predicted egg salad preferences, and then the relationship among implicit belief, explicit belief and recollection.

Experiment 1b

Method

Participants

We collected 163 participants' aIAT data from Experiment 1a. One participant's data were not recorded due to technical issues, thus a total of 162 participants' data were analyzed (36 males, 126 females; M_{age} = 22.23, SD = 3.49). We also collected aIAT data of additional 38 participants (5 males, 33 females; M_{age} = 20.10, SD = 2.51) who did not go through the suggestive session, serving as an extra control group.

Materials and Procedure

Autobiographical Implicit Association Test (aIAT). The aIAT tests participants' implicit truth value (i.e., implicit belief) towards a critical event (see Sartori et al., 2008; Shidlovski, Schul, & Mayo, 2014). The aIAT was programmed using Inquisit 3.

Participants completed five blocks of categorization trials. There were two types of sentences to be categorized: Sentences that were definitely true or false at the time of the experiment and sentences regarding the critical egg salad event. Participants were asked to categorize the sentence as fast as possible.

As Table 6.3 shows, in Block 1 (20 trials), participants categorized definitely “true” (e.g., “I’m sitting now) or “false” (e.g., “I’m climbing a mountain”) sentences. They pressed the left “A” key if the sentence was true and the right “L” key if it was false. In Block 2 (20 trials), participants classified the egg salad sentences to the critical “sick” (e.g., “I got sick after eating egg-salad”) or “not sick” (e.g., “I never got ill eating egg-salad”) categories. They pressed the left “A” key if the sentence was about getting sick on egg salad and pressed the right “L” key if it was about not getting sick on egg salad. Block 3 was a critical phase in which participants pressed the left key if the sentence was true or “sick” sentences and pressed the right key if the sentence was false or about “not sick”. Block 3 contained 20 practice trials and 40 test trials. In Block 4, participants pressed the left key for “not sick” sentences and the right key for “sick” sentences, which was the reverse of Block 2. Block 5 was the other critical phase that was the reversal of Block 3. Participants pressed the left key for true and “not sick” sentences and the right key for false and “sick’ sentences. If RTs were faster in Block 3 than that in Block 5, “sick” sentences were assumed to be implicitly true for the respondent; whereas if RTs were faster in Block 5, “not sick” sentences were true for that respondent.

Table 6.3

Blocks of the Autobiographical Implicit Association Test

Response key	Block 1: logical discrimination	Block 2: Initial autobiographical discrimination	Block 3: Initial double categorization	Block 4: reversed autobiographical discrimination	Block 5: reversed double categorization
“A” key	True sentences	“sick on egg-salad” sentences	True and “sick” sentences	“not sick on egg-salad” sentences	True and “not sick” sentences
“L” key	False sentences	“not sick on egg-salad” sentences	False and “not sick” sentences	“sick on egg-salad” sentences	False and “sick” sentences

Note: The order of Blocks 3 and 5, and of Blocks 2 and 4, were reversed for half of the participants. Sentences in each block were presented in a randomized order.

Table 6.4

Sentences used in the aIAT task

Category	Sentences
True for everyone	I'm sitting now I'm in a little room with a computer I'm doing a psychology experiment I'm in the psychology laboratory I'm in front of the computer
False for everyone	I'm climbing a mountain I'm at the beach I'm eating in a downtown restaurant I'm playing football I'm in a shop
Sick on egg salad	I got sick after eating egg-salad I became ill eating egg-salad when I was little I felt very sick after I ate egg-salad I felt nauseous in my stomach after eating egg-salad I once went to see a doctor because of eating egg-salad
Not sick on egg salad	I did not get sick after eating egg-salad I never got ill eating egg-salad I never felt sick after I ate egg-salad I did not feel nauseous after eating egg-salad I never went to see a doctor because of eating egg-salad

Note: The “True for everyone” and “False for everyone” sentences were from Sartori et al. (2008). The sick or not sick on egg salad sentences have strictly mimicked the guilty and innocent sentences of Sartori et al. (2008).

Results

According to the formula in Greenwald, Nosek, and Banaji (2003), RTs shorter than 400 ms or longer than 10,000 ms were first discarded. Out of 19684 responses, two were under 400ms and 24 were longer than 10,000 ms, thus 0.13% of the responses were deleted. Before analyzing the data, D scores were computed for each participant. D score was calculated by subtracting the mean RT for the block associating “sick” and true sentences (Block 3, including both practice and test trials) from mean RT for the block associating “not sick” and true sentences (Block 5) and then dividing the difference by the inclusive standard deviation of the two blocks. RTs of error responses were included as a latency penalty for errors (Greenwald et al., 2003; Sartori et al., 2008). A positive D score indicates that "sick on egg salad" was implicitly true for participants while a negative D score indicates that it was implicitly not true for participants.

Implicit/explicit Belief and Food Preferences

Participants were categorized in the "Yes" or "No" Implicit Belief categories based on their D scores. For participants who had positive D values, they were considered to hold an implicit belief in the event of sick on egg salad; and for participants who had negative values, they were considered not having an implicit belief of the critical event (Agosta & Sartori, 2013). In total, there were 36 participants who had no implicit belief and 126 participants with implicit beliefs. We tested whether implicit belief of the critical food event would impact people's egg salad preferences. The Mann-Whitney U Test showed that participants with implicit beliefs (M = 3.97, 95%CI [3.57, 4.37]) did not differ from participants without implicit beliefs (M = 4.33, 95%CI [3.61, 5.06]) in preference scores over egg salad, p = .35. Correlation analysis indicated no correlation between D scores and preference scores over egg salad, r = .03, p = .71, with baseline preference scores in Session 1 controlled. There was no correlation either between D scores and egg salad eating intention, r = -.03, p = .68, and no correlation between D scores and preferences for boiled eggs, r = -.06, p = .46.

Since explicit belief was known to impact food preferences in Experiment 1a, we specifically conducted a 2 (Explicit belief: yes vs. no) × 2 (Implicit belief: yes vs. no) ANCOVA (controlling for egg salad preferences in Session 1) to test whether implicit belief interacted with explicit belief in influencing egg salad preferences. No significant interaction effect was detected, F(1, 157) = 0.05, p = .83. There was only a significant main effect of explicit belief, F(1, 157) = 5.03, p = .03, and no main effect of implicit belief, F(1, 157) = 0.01, p = .94. The results overall suggest that implicit belief in the event of being sick on egg salad does not lead to decreased preferences in egg salad.

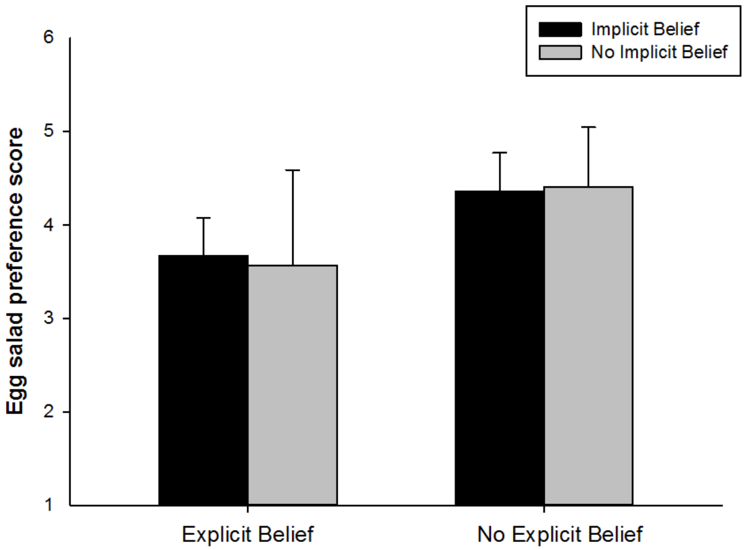


Figure 6.3. Preference scores over egg salad in different explicit belief (yes vs. no) and implicit belief (yes vs. no) groups. Error bars 95%CI.

Explicit Belief, Recollection and Implicit Belief

Implicit belief was measured by the D score in aIAT, while explicit belief and recollection were self-reported in the belief and recollection questions of the AMBQ. We conducted a 2 (Recollection: yes vs. no) × 2 (Explicit Belief: yes vs. no) ANOVA on the D scores, to examine whether D scores were sensitive to dichotomous Recollection and Belief. The assumption of equal variance was met, Levene's test $F(3, 158) = 1.35$, $p = .26$. There was a statistically significant main effect of Recollection, $F(1, 158) = 4.87$, $p = .03$, Cohen's $d = 0.54$, and also a significant main effect of Belief, $F(1, 158) = 4.23$, $p = .04$, Cohen's $d = 0.52$. No significant interaction between Recollection and Belief was found, $F(1, 158) = 0.45$, $p = .51$, see Figure 6.4.

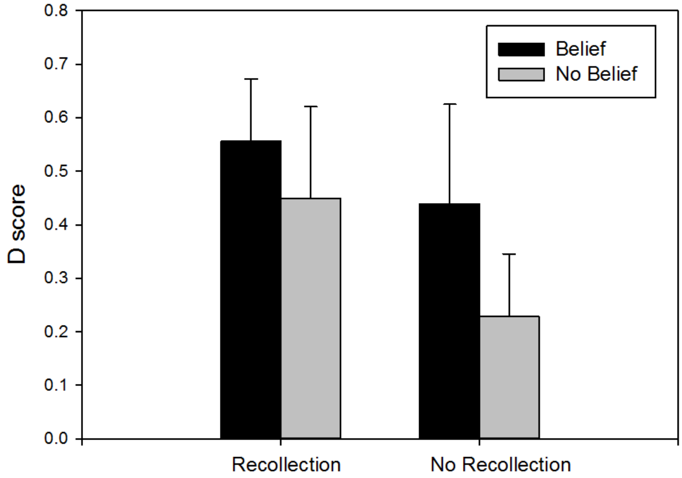


Figure 6.4. Mean D scores in different recollection and belief groups (error bars 95%CI).

As we recorded recollection and belief strength (1 to 8 ratings) in the critical event, we tested the correlations between those ratings and the D scores. Results were similar as in the ANOVA analysis. D scores were correlated significantly with both recollection ratings, $r_{\text{recollection-D score}}(162) = .20$, $p = .01$, and belief ratings, $r_{\text{belief-D score}}(162) = .27$, $p = .001$. The results showed that the more vivid recollections participants experienced, the higher D scores they would have, and the more they believed in the occurrence of the event, the higher D scores in the aIAT.

Relation between Phenomenology and Implicit Belief

In Session 2, all participants went through the guided imagery which was used to evoke rich false autobiographical memories and then their memory characteristics of the event were measured by the MCQ. We were interested in whether the phenomenological characteristics of the suggestive event were related to the D scores. First, we compared phenomenological scores of participants who had positive implicit truth value with participants who had negative implicit truth value. As Table 6.5 shows, participants with implicit belief had higher ratings in every memory characteristic measured, including visual details, smell, taste, setting, time, negative emotion and coherent story (*p*s < .05) than people with no implicit belief in the suggested event.

Table 6.5
Mean phenomenological scores with 95%CI in different implicit truth (yes vs. no) groups

		Visual details	Smell	Taste	Setting
Implicit Truth	Yes (n=126)	4.46 [4.15, 4.77]	4.61 [4.28, 4.95]	4.62 [4.30, 4.94]	4.87 [4.52, 5.22]
	No (n=36)	3.36 [2.69, 4.03]	3.11 [2.37, 3.85]	3.56 [2.87, 4.24]	3.89 [3.07, 4.71]
p value (Mann-Whitney)		.005	<.001	.005	.027
Cohen's d		0.67	0.72	0.65	0.46

Table 6.5 (*Continued*)

		Time	Negative emotion	Coherent story	Average
Implicit Truth	Yes (n=126)	3.50 [3.18, 3.82]	4.10 [3.77, 4.42]	3.20 [2.88, 3.52]	4.19 [3.95, 4.43]
	No (n=36)	2.14 [1.66, 2.62]	3.33 [2.72, 3.95]	2.06 [1.52, 2.59]	3.06 [2.55, 3.58]
p value (Mann-Whitney)		<.001	.024	<.001	<.001
Cohen's d		0.86	0.42	0.68	0.79

Second, we conducted a regression analysis using average phenomenological characteristic (i.e., average of all phenomenological aspects) to predict D scores. One outlier was deleted. The model in which phenomenological characteristic was used to predict D score was significant, $F(1, 157) = 16.32$ $p < .001$, $R^2 = 9.4\%$, see Figure 6.5. The above results section shows that D scores were correlated with both recollection and belief ratings, so we put the three predictors—recollection rating, belief rating, and average phenomenological score—in one regression model. Belief ratings and phenomenological scores were significant predictors of D scores, $F(2, 157) = 9.35$, $p < .001$, $R^2 = 10.6\%$.

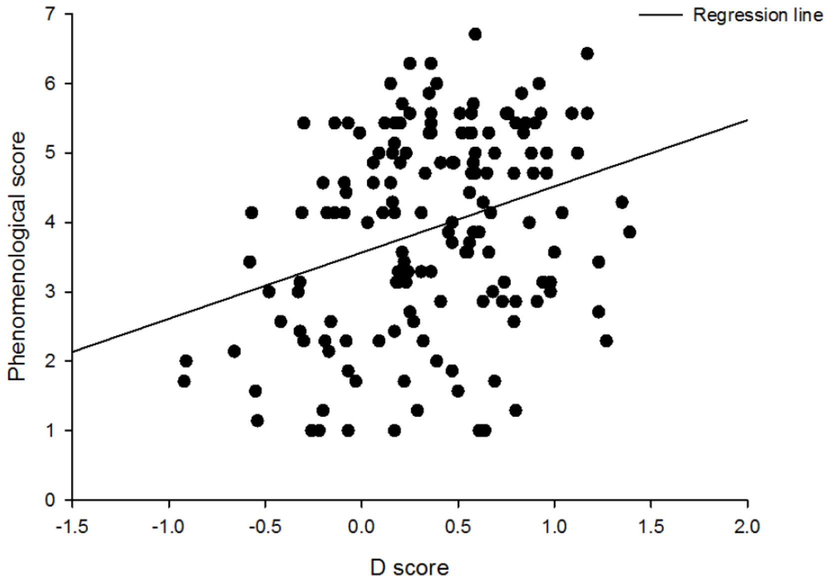


Figure 6.5. Scatter plot of average phenomenological score (Y axis) and D score (X axis) pairs.

Truth Detection Rate of aIAT

Table 6.6 shows the number of participants with or without implicit belief in the critical food event in different recollection/belief groups. We conducted a chi-square test to examine whether the distribution of implicit belief in each recollection/belief group was equal. A chi-square test revealed that participants who had implicit truth value of the critical event tended to be more likely distributed in groups with either a recollection or belief or both, $\chi^2(3) = 9.58$, $p = .02$, $Phi = .24$. As Table 6.6 shows, the aIAT identified more than 82% of participants as holding an implicit truth in those with either a recollection or a belief. However, for participants with neither a recollection nor a belief, still 66.2% of participants were categorized as implicitly encoding the critical food event as authentic. Since all participants went through the suggestion session, we wondered whether suggestion increased false positive rate in the no recollection and no belief group. We therefore collected aIAT data of an additional 38 participants who received no suggestion and had no recollection and no belief of the critical event. In this extra control group, still a high percentage of participants (71%, n = 27) were identified as having implicit truth value by aIAT.

Table 6.6
Distribution of participants with Yes/No Implicit Truth in different recollection/belief groups

		Recollection		No Recollection		Extra Control group
		Belief	No Belief	Belief	No Belief	
Implicit Truth	Yes	40 (88.9%)	18 (85.7%)	23 (82.1%)	45 (66.2%)	27 (71.1%)
	No	5 (11.1%)	3 (14.3%)	5 (17.9%)	23 (33.8%)	11 (28.9%)
Total		45	21	28	68	38

Discussion

We report on the first study examining the implicit belief of false memories and their behavioral consequences. Participants produced false recollections or false beliefs in the event of being sick on egg salad in their childhood. Then they were measured on their implicit beliefs by the aIAT to see whether implicit beliefs about the egg salad event impacted their preferences over egg salad.

First, we found that implicit belief in an egg salad event did not predict preferences for egg salad, but only explicitly self-reported belief in the egg salad event impacted egg salad preference scores. The results support the idea that implicit measurements might not be good predictors of behavior and explicit measurements may be better predictors of behavior than IATs (Oswald et al., 2013, 2015). In previous studies where a correlation between IATs and behavior was found, there was also a correlation between IATs and explicit measurements (see Greenwald et al., 2009). One possible explanation would be that the correlation between IATs and behavior is actually mediated by explicit measurements. Further research is needed to study if there is a possible mediation effect of explicit measurements.

Second, we found that both recollections and beliefs of an event impacted aIAT performance, suggesting the aIAT is sensitive to both components of memory. The results also showed that D scores were positively correlated with recollection strength as well as belief strength. Most intriguingly, the phenomenological characteristics of recollections predicted D scores best, explaining 9.4% of variance of D scores. Participants with implicit beliefs showed significantly higher characteristic scores than those without implicit beliefs on every aspect of memory phenomenology that were measured (e.g., visual details, smell, taste, setting, etc).

Besides, we found the aIAT with good truth detection accuracy (> 82%) in detecting people with either a recollection or a belief, however the false positive rate was very high (66% ~71%). The aIAT tends to detect what is true in participants' mind, although participants' recollections and beliefs were actually false in this case. Even for participants who were debriefed that their memories were false and they had nonbelieved memories, the egg salad event was still labeled as authentic by the aIAT. The results were consistent with previous research showing that mere imagination can greatly impact D scores (Shidlovski, Schul, & Mayo, 2014). Thus it may be not appropriate to use the aIAT in practical settings to distinguish true memories from false memories (see also Takarangi, Strange, & James, 2013).

General Discussion

Memories usually consist of vivid mental representations of a past experience as well as a belief in the event that it did actually occur. In Experiment 1a, we reported if it is recollection or belief of a food aversive event that impacts food preferences. The results showed that belief in the occurrence of the food aversive event, rather than recollective memory traces of the food aversive event, decreased people's food preferences. In Experiment 1b, we reported whether implicit belief measured by aIAT would impact food preferences like explicitly self-reported belief did, and we found that implicit belief did not impact food preferences. We found that aIAT performance was associated with both recollections and beliefs (see Figure 6.6).

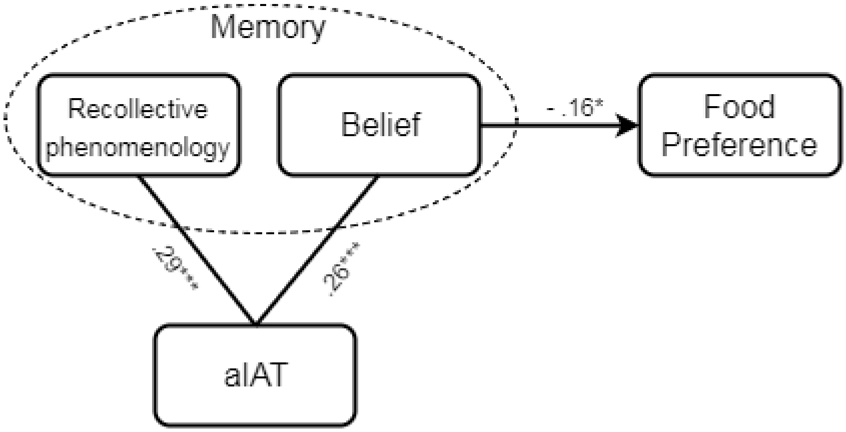


Figure 6.6. Summarized correlations among memory phenomenology, belief, aIAT and food preference in our study (* < .05; *** < .001).

Taken together, the results suggest that food behavior is mostly impacted by a metacognitive judgment of memory instead of memory traces or implicit measurements of memory, which is consistent with Bernstein et al.'s (2015) mega-analyses. In our study, around 41% of the participants reported false recollections of the egg salad event, however those with false recollections did not differ from those with no recollections in egg salad preferences. Meanwhile, people who believed in the egg salad event showed lower egg salad preference than people who had no beliefs. Thus, the consequences of memory are likely to be mediated by autobiographical belief, which plays a central role in remembering. Using structural equation modeling, Scoboria et al. (2014) found that recollection is predicted by perceptual, temporal and spatial aspects of a past experience, while belief is strongly predicted by event plausibility. When a person has a belief in the occurrence of the event, then the event is seen as highly plausible and that it may happen again. In our case, when participants had a belief that they had been sick after eating egg salad in childhood, participants can naturally reason that they would probably get sick again after eating egg salad. Thus, their preferences for egg salad were lower than those who did not have a belief.

If recollection in general cannot impact food behavior, what is the function of recollection? Based on our data, we speculate that recollection can perhaps facilitate information processing efficiency or fluency. The results showed that recollection of the event was closely related to implicit measurement (i.e., reaction times) of the event. The more vivid and detailed of the recollection, the higher D score or faster reaction time to the statements describing the true status of the event (Experiment 1b). The mental representation of the event may have made the processing of related information (e.g., an imagery depicting the event or in our study, the statement that "I have been sick after eating egg salad") more fluent than the no recollective representation condition. Recollection has also been shown to contribute to simulating a future event, which can help people plan and cope with the future (Schacter, 2012).

The idea that recollection facilitates processing fluency is in line with research on the consequences of inaccurate information. Research has found that people can still be impacted by the content of the false information (e.g., Newton proposed the theory of relativity) even when they know the information is false (Fazio, Barber, Rajaram, & Ornstein, 2013; Rapp, 2016). Researchers have used the fluency model, which emphasizes the experience of processing information, to explain the impact from recollections of false information (Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012). For instance, if people read a false statement that "The Sun revolves around the earth" which they know is false, and when they are asked "____revolves around the earth" later, "The Sun" will pop up automatically in mind since it contributes to a fluent flow of message. This is similar to our study that participants categorized sentences like "I got very sick after eating egg salad" faster into true categories than into false categories when they had mental images of being sick after eating egg salad, and also when they believed the statement.

Our results demonstrate an interesting alignment between the model of (explicit) autobiographical memory constructs (Scoboria et al., 2014) and implicit measurement of autobiographical events (Agosta & Sartori, 2013). In the two-dimensional model of explicit memory, recollection and autobiographical belief are two distinct constructs. The current study managed to produce all possible recollection (yes vs. no) and belief (yes vs. no) combinations and then their relationship with the aIAT was tested. Our results showed that the aIAT was sensitive to detect both recollection and belief change. An ANOVA analysis as well as correlational analyses consistently showed that the stronger recollection people had, the higher D score of aIAT, and the stronger belief people had, the higher the D score (see Figure 6.6).

One might suspect, if the aIAT is sensitive to detect both recollection and belief, why would it not predict food preferences since belief impacts food preferences? The answer may lie in the different contribution of recollection and belief in explaining the variance of aIAT D scores. The results showed that memory phenomenological characteristics contributed the most to D scores, explaining around 9.4% of the variance, while by adding belief in the regression model, the overall R^2 increased slightly to 10.6%. Thus the aIAT is much more sensitive to recollective aspects of memory than to belief judgments of a memory. And since recollection does not predict food preferences, as a result, aIAT probably cannot predict food preferences. Then in situations where implicit measurements have strong correlations with explicit measurements, implicit measurements might be correlated with behavior as well. This can explain results from previous studies where implicit attitude to black people were found to be correlated with discriminatory behavior since IATs were also correlated with explicit attitude towards black people (e.g., McConnell & Leibold, 2001; Greenwald et al., 2009).

Marini, Agosta, Mazzoni, Barba, and Sartori (2012) investigated aIATs of true and false memories of words using the Deese/Roediger-McDermott (DRM) paradigm. They found that the aIAT detected greater D scores for true memories than false memories. However, average D scores for both true and false memories for words were positive (0.98 and 0.86 respectively) in their study. Our study has similar results in that the average D scores for all (false) recollection and belief groups were positive (0.23~0.56), suggesting that the aIAT is limited in differentiating true from false events when recollections or beliefs for the false event have developed. In our study, we also found a very high false positive rate (around 70%) for participants who had neither false beliefs nor false recollections. We could not find an explanation for it since the sentences in our study were constructed in a very similar way with previous research (Sartori et al., 2008). Further research is needed to study the truth detection validity of aIAT.

There are certain limitations of the current study. First, the categorization of participants into different groups was based on self-reported recollections and beliefs instead of being randomly assigned. Participants were naturally assigned to different recollection and belief groups based on their self-reports at the end of Session 2, since we could not control which participants to develop a false belief/recollection and which not. Thus the groups in ANOVA analyses of the current study were not based on manipulations but based on the output of our manipulations. Thus, the sample size of the groups varied to some extent. For instance, in the nonbelieved memory group, there were 21 participants while in the believed memory group, there were 69 participants. Again, this is because we cannot control which participants would relinquish their beliefs after the debriefing manipulation. Further research may try to increase the number of nonbelieved memories by providing more persuading feedback, in order to even the number of participants. To overcome this limitation to some extent, we conducted a regression analysis with dummy variables of recollection (yes vs. no) and belief (yes vs. no), and the results were similar.

The other limitation is that the memories studied here are false memories. Participants produced false recollections and false beliefs of the egg salad event after a suggestion session. One might suspect that, because those memories were essentially false, it is relatively easy for participants to relinquish beliefs in their recollections. The role that belief plays in false memories might be different from that it plays in true memories. However previous studies showed that autobiographical belief in true memories, such as performing actions and studying words, is retracted as easily as belief in false memories via social feedback (Mazzoni, Clark, & Nash, 2014; Wang et al., 2017). Nonetheless, further research is needed to study whether the conclusion that only autobiographical belief impacts food preferences can be generalized to true memories. Further research is also needed to study the consequences of recollection and belief with other behavioral measurements.

Our study has theoretical implications regarding the dissociation theory of memory. Our study managed to create all possible dichotomous recollection and belief combinations, which supported the dissociation model of these two components in memory. Nonbelieved memories are relatively rare in real life (3 ~ 6.4 %, Scoboria & Talarico, 2013), but in lab settings, a relatively higher rate (13%) of NBMs can be created. Thus the different roles of recollection and belief could be tested in our study, and the current study found dissociation in the consequences of recollection and belief. We have provided evidence that autobiographical belief is important in impacting food behavior while recollection impacts the processing of event-related information.

The current study also has practical implications in legal psychology. In the legal arena, eyewitness misidentification happens much more frequently than other factors contributing to miscarriages of justice (innocenceproject.org). Our study supports the view that autobiographical belief is the key in impacting attitudes and behavior (Bernstein et al., 2015). Under the legal context, the results have raised the concern that to what extent eyewitness misidentification is impacted by false recollection and/or false belief. For instance, Smeets, Telgen, Ost, Jelicic, and Merckelbach's (2009) study on false memories showed that people sometimes claimed to have memories of seeing footage of an airplane crash while in fact they only had (false) beliefs in seeing the footage. Moreover, autobiographical belief is even more malleable than recollection (Scoboria et al., 2014). Varied sources of information such as social feedback from the police, one's own motivation, the media, views of other people, etc., can all possibly change people's autobiographical beliefs in remembering past events (Scoboria, Boucher, & Mazzoni, 2014). Further research is needed to examine the contribution of the components of eyewitness memories, i.e., recollection and belief, in impacting their reports or identifications.

References

Agosta, S., & Sartori, G. (2013). The autobiographical IAT: a review. *Frontiers in Psychology*, *4*, 519.

Arcuri, L., Castelli, L., Galdi, S., Zogmaister, C., & Amadori, A. (2008). Predicting the vote: Implicit attitudes as predictors of the future behavior of decided and undecided voters. *Political Psychology*, *29*, 369-387.

Berkowitz, S. R., Laney, C., Morris, E. K., Garry, M., & Loftus, E. F. (2008). Pluto behaving badly: False beliefs and their consequences. *The American Journal of Psychology, 121*, 643-660.

Bernstein, D. M., & Loftus, E. F. (2009). The consequences of false memories for food preferences and choices. *Perspectives on Psychological Science*, *4*, 135-139.

Bernstein, D. M., Laney, C., Morris, E. K., & Loftus, E. F. (2005). False beliefs about fattening foods can have healthy consequences. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 13724-13731.

Bernstein, D. M., Scoboria, A., & Arnold, R. (2015). The consequences of suggesting false childhood food events. *Acta Psychologica*, *156*, 1-7.

Clark, A., Nash, R. A., Fincham, G., & Mazzoni, G. (2012). Creating non-believed memories for recent autobiographical events. *PLOS One*, *7*, 1-7.

Fazio, L. K., Barber, S. J., Rajaram, S., Ornstein, P. A., & Marsh, E. J. (2013). Creating illusions of knowledge: Learning errors that contradict prior knowledge. *Journal of Experimental Psychology: General*, *142*, 1-5.

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

Greenwald, A. G., Banaji, M. R. & Nosek, B. A. (2015) Statistically small effects of the Implicit Association Test can have societally large effects. *Journal of Personality and Social Psychology*, *108*, 553–61.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*, 197-216.

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the implicit association test. *Journal of Personality and Social Psychology*, *97*, 17-41.

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.

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.

Lewandowsky, S., Ecker, U. K., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, *13*, 106-131.

Loftus, E. F. (1993). The reality of repressed memories. *American psychologist*, *48*, 518-537

Mantonakis, A., Wudarzewski, A., Bernstein, D. M., Clifasefi, S. L., & Loftus, E. F. (2013). False beliefs can shape current consumption. *Psychology*, *4*, 302-308.

Marini, M., Agosta, S., Mazzoni, G., Dalla Barba, G., & Sartori, G. (2012). True and false DRM memories: differences detected with an implicit task. *Frontiers in psychology*, *3*, 310.

Mazzoni, G. A., Clark, A., & Nash, R. A. (2014). Disowned recollections: Denying true experiences undermines belief in occurrence but not judgments of remembering. *Acta Psychologica*, *145*, 139-146.

Mazzoni, G. A., 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.

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

McConnell, A. R., & Leibold, J. M. (2001). Relations among the Implicit Association Test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology*, *37*, 435-442.

Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. (2013). Predicting ethnic and racial discrimination: A meta-analysis of IAT criterion studies. *Journal of personality and social psychology*, *105*, 171-192.

Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. (2015). Using the IAT to predict ethnic and racial discrimination: Small effect sizes of unknown societal significance. *Journal of Personality and Social Psychology, 108*, 562–571.

Otgaar, H., Scoboria, A., & Mazzoni, G. (2014). On the existence and implications of nonbelieved memories. *Current Directions in Psychological Science, 23,* 349-354.

Otgaar, H., Scoboria, A., & Smeets, T. (2013). Experimentally evoking nonbelieved memories for childhood events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 717-730.

Otgaar, H., Scoboria, A., Howe, M. L., Moldoveanu, G., & Smeets, T. (2016). Challenging memories in children and adults using an imagination inflation procedure. *Psychology of Consciousness: Theory, Research, and Practice*, *3*, 270-283.

Rapp, D. N. (2016). The consequences of reading inaccurate information. *Current Directions in Psychological Science*, *25*, 281-285.

Rubin, D. C. (2006). The basic-systems model of episodic memory. *Perspectives on Psychological Science, 1,* 277-311.

Sartori, G., Agosta, S., Zogmaister, C., Ferrara, S. D., & Castiello, U. (2008). How to accurately detect autobiographical events. *Psychological Science*, *19*, 772-780.

Schacter, D. L. (2012). Adaptive constructive processes and the future of memory. *American Psychologist*, *67*, 603-613.

Scoboria, A., Boucher, C., & Mazzoni, G. (2015). Reasons for withdrawing belief in vivid autobiographical memories. *Memory, 23,* 545–562.

Scoboria, A., Jackson, D., Talarico, J., Hanczakowski, M., Wysman, L., & Mazzoni, G. (2014). The role of belief in occurrence within autobiographical memory. *Journal of Experimental Psychology: General, 143,* 1242-1258.

Scoboria, A., Mazzoni , G. , Jarry, J. , & Bernstein, D. (2012). Personalized and not general suggestion produces false autobiographical memories and suggestion-consistent behavior . *Acta Psychologica*, *139*, 225–232 .

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

Scoboria, A., Talarico, J. M., & Pascal, L. (2015). Metamemory appraisals in autobiographical event recall. *Cognition, 136,* 337-349.

Shidlovski, D., Schul, Y., & Mayo, R. (2014). If I imagine it, then it happened: the implicit truth value of imaginary representations. *Cognition*, *133*, 517-529.

Smeets, T., Telgen, S., Ost, J., Jelicic, M., & Merckelbach, H. (2009). What's behind crashing memories? Plausibility, belief and memory in reports of having seen non-existent images. *Applied Cognitive Psychology*, *23*, 1333-1341.

Takarangi, M. K., Strange, D., Shortland, A. E., & James, H. E. (2013). Source confusion influences the effectiveness of the autobiographical IAT. *Psychonomic Bulletin & Review*, *20*, 1232-1238.

Wang, J., Otgaar, H., Howe, M. L., Smeets, T., Merckelbach, H., & Nahouli, Z. (2017). Undermining belief in false memories leads to less efficient problem-solving behavior. *Memory*, *25*, 910-921.

Part III

Theoretical Accounts and General Discussion

CHAPTER 7

Memory as an Adaptive, Reconstructive System Related to the Self

This Chapter is an adapted and combined version of the following papers:

Wang, J., Otgaar, H., Howe, M. L., & Zhou, C. (in press). A self-reference false memory effect in the DRM paradigm: Evidence from Eastern and Western samples. *Memory & Cognition*.

Zhou, C., Wang, J., & Zhou, W. (2014). The self-reference effect in false memory: A robust facilitation of self. *Journal of Psychological Science, 5*, 1079-1083. (in Chinese)

Abstract

It is well established that processing information in relation to oneself (i.e., self-referencing) leads to better memory for that information than processing that same information in relation to others (i.e., other-referencing). However, it is unknown whether self-referencing also leads to more false memories than other-referencing. In the current two experiments with European and East Asian samples, we presented participants with the Deese-Roediger/McDermott (DRM) lists together with their own name or other people's name (e.g., Trump). We found consistent results across the two experiments; that is, in the self-reference condition, participants had higher true and false memory rates compared to those in the other-reference condition. Moreover, we found that self-referencing increased recollection scores of both true and false memories but not familiarity of the items using Remember/Know judgements. These findings are discussed in terms of theoretical frameworks such as spreading activation theories and the fuzzy trace theory. We propose that our results reflect the adaptive nature of memory in the sense that cognitive processes that increase mnemonic efficiency may also increase susceptibility to associative false memories.

Keywords: Self-reference, false memory, recollection, spreading activation, adaptive memory

The self plays an important role in memory (Conway, 2005; Howe, 2014; Klein, 2012). It is well established that information is better remembered when it is processed in relation to the self than when it is processed in relation to other sources (Symons & Johnson, 1997). For instance, judging words for their personal relevance (“Does this word describe you?”) produces superior memory recall than other-referencing (“Does this word describe Henry?”) or semantic encoding (“Does this word mean the same as XXX?) (Rogers, Kuiper, & Kirker, 1977; Kuiper & Rogers, 1979). The superior mnemonic effect for items encoded with respect to the self, dubbed the self-reference effect (SRE), has received considerable empirical support. For instance, SRE has been found in various age groups (from children to old people; Cunningham, Brebner, Quinn, & Turk, 2014; Gutchess, Kensinger, Yoon, & Schacter, 2007) as well as in different cultures (Zhu, Zhang, Fan, & Han, 2007).

One of the proposed mechanisms underlying SRE is relational processing or organizational processing (Klein, 2012; Klein & Loftus 1988; Symons & Johnson, 1997). Relational processing refers to the encoding of associations among items in a list leading participants to think about the shared concept or label (Klein, 2012; Klein & Loftus, 1988). For example, when participants see the words “sound” and “piano”, the items can be organized as concepts related to “music”. Another example would be items sharing the same category, such as “dog” and “horse” that can be both processed under the category “animal”. Klein and Loftus (1998) manipulated the relatedness among items within a list and found that for weakly related lists, category sorting (i.e., relational processing) produced equivalent recall as self-referenced processing, suggesting that relational processing is one of the mechanisms underlying SRE.

In line with this, Sui and Humphreys (2015) have recently proposed that self-reference increases the binding among different forms of information in memory. For instance, evidence shows that people not only remember the self-referenced stimuli better than other-referenced stimuli, but also remember the stimuli−self associations (e.g., an apple paired with one’s own face) better than the stimuli−other associations (e.g., a cup paired with other people’s face) (Cunningham et al., 2014). Leshikar and Duarte (2014) reported that both young and old adults exhibited better relational memories for object-scene bindings (e.g., saxophone-beach) in the self-reference rather than the other-reference condition. According to Sui and Humphreys, the self acts as a form of associative glue in a memory network and whenever self-representations are activated, they strengthen the binding among different stimuli in relation to the self-representation, thus increasing retention for experienced stimuli. Taken together, research so far has demonstrated that relational processing among the list items or the binding among stimuli may account for the mnemonic superiority of self-referenced processing.

Interestingly, according to spreading activation theories of false memory (e.g., Activation Association Theory, AAT, Howe, Wimmer, Gagnon, & Plumpton, 2009; Activation/Monitoring Theory, AMT, Roediger, Balota, & Watson, 2001), relational processing or binding between items may also contribute to the creation of false memories. False memories refer to memory for details or events that were not experienced. The Deese-Roediger/McDermott paradigm (DRM; Deese, 1959; Roediger & McDermott, 1995) is a typical experimental paradigm used to study false memories. Here, participants study lists of associated words (e.g., *sound, piano, sing, radio, band, melody, horn, concert, instrument*), and during memory tests, participants oftentimes falsely recall or recognize a non-presented critical lure (i.e., *music*) that is associated with the list items as having been presented to them. Studies with the DRM paradigm show that increasing the associative strength of a studied list (i.e., increasing relational processing) results in higher levels of false memories (Cann, McRae & Katz, 2011; Gallo & Roediger, 2002; Howe et al., 2009). Following this line of reasoning, enhancing relational processing using self-referencing should not only make people better remember what has been presented, but should also enhance the susceptibility to the DRM memory illusion. The current experiments were conducted to examine this idea.

Spreading activation theories (e.g., Howe et al., 2009) suggest that related concepts, such as the DRM list items, are embedded within an associative network and that the presence of an item (e.g., the word “piano”) activates a corresponding concept (see also Otgaar, Muris, Howe, & Merckelbach, in press). Most importantly, according to these theories, activation will automatically spread to “nearby” related concepts (e.g., *music*) and the level of activation determines the rate of recall/recognition (Anderson, 1983). The activation of related but non-presented concepts leads to false recollections of non-presented critical lures. As Figure 7.1 shows, we hypothesize that, when the DRM list items are processed in relation to the self, the self may strengthen the relatedness among the items as well as the activation spreading to the critical lures, thus evoking more false memories of the critical lures than when relating information to others.

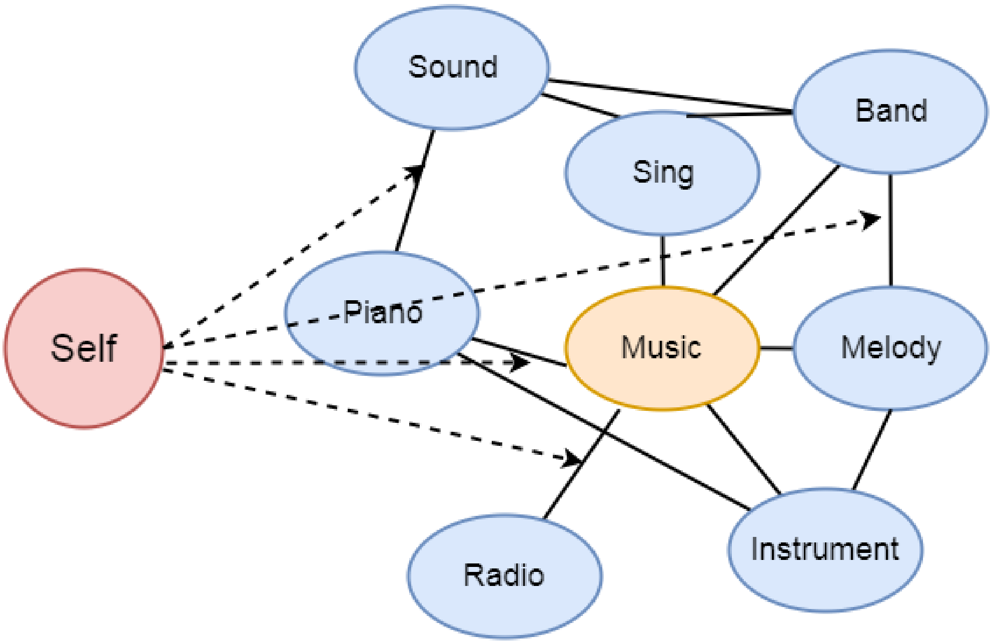


Figure 7.1. Hypothetical associative memory network when the self is referenced, based on spreading activation theories (Howe et al., 2009; Roediger et al., 2001) and the mechanisms of SRE (Klein, 2012; Sui & Humphreys, 2015). Blue circles represent DRM list items; yellow circle represents the critical lure; solid lines represent relations among items and dashed arrows represent the strengthening effect to the relatedness among items.

Indeed, there is some evidence suggesting that self-reference may be related to false memories. Rogers, Rogers, and Kuiper (1979) found that false alarms for personal adjectives increased as the degree of self-reference increased. Rosa and Gutchess (2013) extended this effect to older adults. They asked participants to rate adjectives (e.g., cultured, sensible) for self-descriptiveness on a 9-point scale and later measured their memories for the adjectives and some adjective lures. The authors found that high self-descriptiveness words led to higher false alarms than low self-descriptiveness words. One limitation with the aforementioned research is that they did not compare the false alarms between the self-reference and other-reference conditions, but only examined the relationship between the degree of self-reference and false alarms. Furthermore, previous studies did not use a well-established false memory paradigm (e.g., the DRM paradigm) known to elicit associative false memories that might be caused by similar mechanisms as the SRE. Choosing such a paradigm would enable us to test the mechanisms underlying SRE and false memories.

To date, no study has compared the false memories of self- vs. other-reference conditions and no study has examined the relationship between true and false memories (i.e., net accuracy) generated from self-referential processing. Net accuracy measures the overall memory accuracy, which is the ratio of true memories to true memories plus false memories (Brainerd, Reyna, & Ceci, 2008). Since relational processing may amplify both true memories and false memories, it is unknown whether self-referencing would increase or rather decrease the net accuracy (i.e., true memory to true memory plus false memory) relative to other-referencing. However, since self-referencing has been demonstrated to be a superior memory encoding strategy in previous literature, one would expect that self-referencing should increase the overall memory accuracy or net accuracy. The present study aimed to examine self-referenced false and true memories as well as net accuracy in the DRM paradigm.

The Current Experiments

We presented participants with DRM lists appearing together with either their own name (self-referenced condition) or some other people's name (i.e., Trump; other-referenced condition). During the study phase, participants were asked to remember the words shown to them as well as which source (self or other) the words appeared together with. After a filler task, they were asked to recognize the words as presented or not in the recognition phase. We predicted that the self-referenced condition would result in higher false recognition rates of the critical lures than the other-referenced condition. We also predicted that self-referencing would lead to the highest true recognition rate of the studied items (i.e., the SRE).

Importantly, previous research has shown that self-reference can increase memory for perceptual details of the studied items (Hamami, Serbun, & Gutchess, 2011; Leshikar, Dulas, & Duarte, 2015; Serbun, Shih, & Gutchess, 2011). Serbun et al. (2011) asked participants to encode objects in relation to self or other and found that self-referenced objects were remembered with more visual details. Leshikar and collegues (2015) used the Remember/Know procedure (R/K) to measure recollection and familiarity of self-referential memories. Recollection and familiarity are two sub-types of recognition memory that can be calculated from R/K responses. Recollection indicates remembering episodic details, while familiarity refers to the feeling of recognizing an item but yet one cannot recollect specific details (Yonelinas, 2002). Leshikar et al. (2015) found that the self-referencing condition exhibited better recollection of studied items than semantic encoding (i.e., processing the meaning of presented words). They also found more phenomenological details on the Memory Characteristic Questionnaire (MCQ) measurements. It is unknown whether self-referencing would also increase the phenomenology of false memories. It might be that self-referencing impacts familiarity of falsely recalled items as Rosa and Gutchess (2013) suggested. However, it is also possible that self-referencing impacts true and false memories in a similar way by increasing recollection as they share the same activation network. We used the R/K judgment procedure (Yonelinas, 2002; Yonelinas & Jacoby, 1995) in the recognition phase to examine the recollection and familiarity of critical lures in self- vs. other-reference conditions.

Experiment 1

Method

Participants

With an estimation of medium effect size (*d* = 0.5; Symons & Johnson, 1997) and power as 0.8, a priori power analysis using G. Power 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007) showed that thirty-four participants were needed. Thirty-nine participants from Maastricht University, the Netherlands were tested. The sample consisted of 8 males and 31 females (M_{age}= 21.6, *SD*= 2.29, min 18 – max 31 years old). All participants were fluent in English. The experiment was approved by the ethical board of Faculty of Psychology and Neuroscience, Maastricht University.

Materials[11]

Eighteen DRM word lists with 12 words per list were used. The lists have been used in previous work as well (e.g., Howe, Garner, & Patel, 2013). DRM lists were pseudo-randomly assigned to the self-reference (6 lists, *Mean* BAS = 0.26, *SD* = .13), other-reference (6 lists, *Mean* BAS = 0.26, *SD* = .12), and control conditions (6 lists, *Mean* BAS = 0.26, *SD* = .10), with the backward association strength (BAS) matched across three

[11] All materials can be accessed from https://osf.io/snej6/

conditions, $F(2, 16) = 0.006$, $p = .995$. BAS is the average association strength between DRM list members and the critical lure. Besides controlling the BAS across conditions, we switched the lists in the self-reference and other-reference conditions for half of the participants (n =19) to eliminate any list effects. The recognition list included 18 critical lures (one critical lure per list), 36 studied items and 36 unrelated new items. The studied items were from the 3rd and 10th position of each list.

Design and Procedure

The experiment employed a 3 (Reference: self vs. other vs. neutral) × 2 (Memory type: true vs. false) within-subject design. All participants were tested individually in a quiet and isolated room. The experiment was programmed using Visual Basic.

In the study phase, participants first were asked to provide some background information, including their own name (the name that they were most frequently referred to). Then the DRM words appeared one by one together with either their own name (self-reference condition) or the name “Trump” (other-reference condition), or the DRM words appeared with no other name next to it (neutral condition). “Trump” was used in the other-reference condition as it was a very familiar name due to the American election at the time of our data collection. Participants were asked to remember the words and to which source they appeared together with (self, Trump, or alone). The DRM words were shown list by list, and the items within a list appeared with the same source (see Figure 7.2). The experiment was separated in 6 blocks with each block containing three BAS matched lists (a self-referenced list, an other-referenced list, and a control or neutral list). The sequence of the lists within a block was randomized.

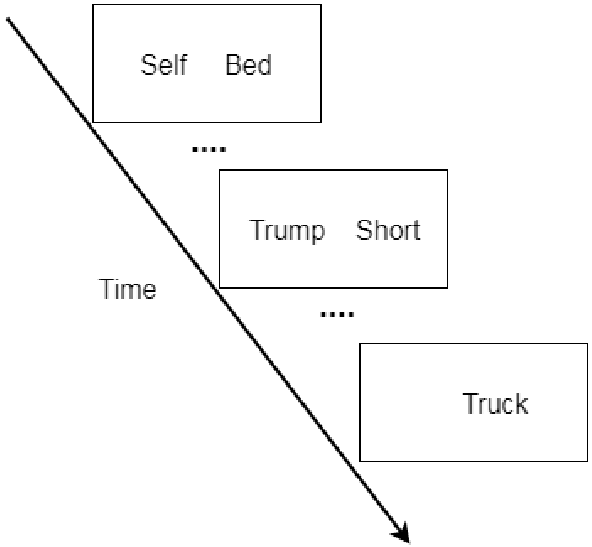


Figure 7.2. Illustration of one block in the study phase ("Self" stands for participant's own name). *Bed*, *Short* and *Truck* are representative words from different lists. Each word pair was presented for 1500 ms with 500 ms inter-stimulus interval.

After the study phase, participants performed a filler task (playing the "Bejeweled" game) for around 5 min. Then, their memories were tested using a recognition task. Words were shown at the center of the screen and participants responded by clicking the "Remember", "Know", or "New" button. If the word was new, they clicked the "New' button; if they remembered the word as old and can recall specific details such as font, size, etc., they clicked the "Remember" button; and if they identified the word as old but could not recall specific details, they clicked the "Know" button (Yonelinas & Jacoby, 1995).

Results[12] and Discussion

Recognition Rates

"Remember" and "Know" responses were combined when calculating the overall "old" recognition response rate. A 3(Reference: self vs. other vs. neutral) × 2 (Memory type: true vs. false) repeated measures ANOVA was conducted on these old recognition rates. We found a statistically significant main effect of Reference, F (2, 76) = 6.77, p = .002, partial

[12] All data can be accessed from https://osf.io/snej6/

η^2 = 0.15. Pairwise comparisons with Bonferroni corrections indicated that the self-reference condition had higher true and false recognition rates than the other-reference condition ($p < .001$) and the neutral condition did not differ from either condition (ps > .10). There was no statistically significant main effect of Memory type, F (2, 76) = 0.22, p = .65, indicating equivalent level of true memories and false memories were recognized. No statistically significant interaction effect between Reference and Memory type was found, F (2, 76) = 0.21, p = .81, suggesting that the self-reference effect is similar for both true and false memories (see Figure 7.3a).

We were specifically interested in comparing the effect sizes of self-reference effect in false memories and true memories, respectively. The result showed that false recognition rates in the self-reference condition was significantly higher than the other-reference condition, $t(38)$ = 3.68, p = .001, Cohen's d =0.60. A paired-samples t-test also showed significantly higher true memory rates in self-reference condition than in the other-reference condition, t (38)= 2.02, p = .05, Cohen's d = 0.33.

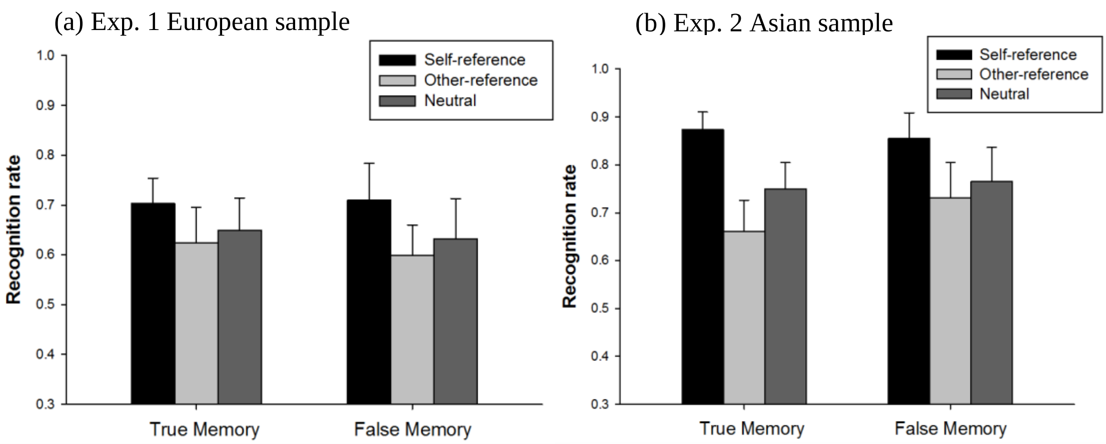


Figure 7.3. True and false recognition rates in self-reference, other-reference and neutral conditions in Experiment 1 (European sample; n= 39) and Experiment 2 (East Asian sample; n= 29) (Error bars 95%CIs).

Net Accuracy

Net accuracy was calculated as the ratio of true recognition to true recognition plus false recognition (Brainerd, Reyna, & Ceci, 2008). We conducted a repeated measures ANOVA to compare net accuracy in the self-reference, other-reference and neutral conditions. No significant main effect was found on the Reference variable, F (2, 76) =

0.10, p = .91. Net accuracy of the self-reference processing (M = 0.51, 95%CI [0.47, 0.54]) did not differ from that of the other-reference processing (M = 0.50, 95%CI [0.46, 0.55]) or the neutral condition (M = 0.51, 95%CI [0.48, 0.55]). The results suggest that self-referencing did not increase the overall accuracy of memory compared to other-referencing and the neutral conditions in the DRM paradigm.

Recollection and Familiarity of Self-referenced False Memory

Recollection is the rate of "Remember" responses, whereas familiarity is calculated as the rate of Knowing responses divided by 1 minus the rate of Remember responses (Knowing rate/(1- recollection rate), Jacoby & Yonelinas, 1995; Yonelinas, 2002). We conducted repeated-measures ANOVAs to compare recollection of critical lures in different reference conditions. As Table 7.1 shows, we found that different Reference conditions led to statistically different recollection rates for critical lures, F (2, 76) = 7.24, p = .001, partial η^2 = 0.16. Follow-up Bonferroni contrasts showed that in the self-reference condition, critical lures received statistically higher recollection scores than in the other-reference (t(38) = 3.55, p = .001, Cohen's d =0.58) and neutral conditions (t(38) = 2.97, p = .005, Cohen's d = 0.48). For familiarity, all conditions had equal levels of familiarity, F (2, 76) = 1.75, p = .18. Thus, processing information related to oneself mainly increased false recollections relative to the condition in which information was related to others.

Table 7.1

Recollection and familiarity of critical lures in different reference conditions of Experiments 1 and 2 (Means with 95%CIs)

Reference	Exp. 1 European sample		Exp. 2 East Asian sample	
	Recollection	Familiarity	Recollection	Familiarity
Self-reference	0.49** [0.40, 0.57]	0.44 [0.34, 0.55]	0.72*** [0.63, 0.80]	0.30 [0.14, 0.46]
Other-reference	0.35 [0.28, 0.43]	0.36 [0.27, 0.44]	0.51 [0.43, 0.59]	0.39 [0.25, 0.53]
Neutral	0.35 [0.27, 0.44]	0.45 [0.35, 0.56]	0.55 [0.45, 0.65]	0.41 [0.26, 0.56]

** p = .001, *** p < .001 (self- vs. other-reference comparison)

Recollection and Familiarity of Self-referenced True Memory

We conducted the same repeated measures ANOVAs on recollection and familiarity of true memories. We found a statistical difference in recollection scores for presented items among the reference conditions, F (2, 76) = 4.97, p = .009, partial η^2 = 0.12, while familiarity was not significantly impacted by Reference, F (2, 76) = 0.65, p = .52. Paired-samples t-tests showed that in the self-reference condition, presented items had statistically higher recollection scores than in the other-referenced condition (t(38) = 2.07, p = .045, Cohen's d = 0.33) and the neutral condition (t(38) = 3.20, p = .003, Cohen's d =0.51). The latter two did not differ from each other, t(38) = 0.75, p = .46. The results were consistent with previous studies that found self-reference increased recollections of true memories (Leshikar et al., 2015).

Table 7.2

Recollection and familiarity of studied items in different reference conditions of Experiments 1 and 2 (Means with 95%CIs)

Reference	Exp. 1 European sample		Exp. 2 East Asian sample	
	Recollection	Familiarity	Recollection	Familiarity
Self-reference	0.48* [0.41, 0.56]	0.42 [0.34, 0.49]	0.77*** [0.71, 0.83]	0.38 [0.26, 0.49]
Other-reference	0.39 [0.32, 0.47]	0.40 [0.32, 047]	0.48 [0.40, 0.56]	0.34 [0.24, 0.43]
Neutral	0.37 [0.30, 0.44]	0.45 [0.37, 0.53]	0.57 [0.49, 0.65]	0.38 [0.28, 0.50]

*p < .05, *** p < .001 (self- vs. other-reference comparison)

In Experiment 1, we showed that in line with our hypothesis, relating information to oneself increased both true and false memories compared to relating information to others. Moreover, we found that self-reference specifically increased recollection scores of studied items and critical lures. These may be explained by our hypothetical mechanism of the SRE in the associative memory network (Figure 7.1). Because the self facilitates the relational processing or the binding of DRM list items (Klein, 2012; Sui & Humphreys, 2015), on the one hand, facilitated relatedness increases the phenomenological level of activated recollection of the studied items, as demonstrated by our results and also previous SRE research (e.g., Leshikar et al., 2015; Serbun et al., 2011). On the other hand, the increased activation may automatically spread to critical lures (e.g., AAT, Howe et al., 2009), and increase the activation level of critical lures as well. As a result, the highest proportion of critical lures that were falsely remembered was in the self-reference condition.

Besides relational processing, the other mechanism that may mediate SRE is item-specific processing (or elaboration) (Klein, 2012). Item-specific processing refers to encoding specific information of an item (Hunt & Einstein, 1981; Klein & Loftus, 1988). For example, for “sound” and “piano”, “sound” can be encoded as something people hear while “piano” is an object. However, item-specific processing distinguishes different items and it has been found to reduce false memories (e.g., Arndt & Reder, 2003; Israel & Schacter, 1997; McCabe, Presmanes, Robertson, & Smith, 2004). For instance, when DRM items within a list were presented in different fonts, false memories were reduced compared to when DRM items were presented in the same font (Arndt & Reder, 2003). Our results that self-referencing increased false memories seem not support item-specific processing as the dominant processing in the current paradigm. Moreover, if item-specific processing is the main mechanism of SRE here, one would expect it to increase net accuracy as it should increase true memory but decrease false memory, which was not the case in our results.

Surprisingly, we found that net accuracy of self-referencing did not differ significantly from other-referencing and neutral conditions, suggesting self-referencing does not increase the overall memory accuracy in a paradigm where relational processing would lead to associative memory illusions. The results showed that self-referencing increased true and false memories to the same extent compared to the other conditions, leading to no statistical difference between conditions. These results indicate that relational processing is the dominant processing in SRE with the DRM paradigm. Relational processing results in an increase of activation of both studied items and critical lures that share the same memory network.

Thus, Experiment 1 shows that processing information in relation to the self elevates true *and* false memory levels. One may wonder why we did not find a statistical difference between the self-reference condition and the neutral condition on memories. The reason might be that there was no reference provided in the neutral condition, leaving participants more time to process the studied items alone. However, this will not hinder our conclusion on SRE as SRE here is based on self- versus other-referencing comparisons. Another interesting observation was that net accuracy was not affected by self-referencing. To examine the reliability of these finding, in Experiment 2, we tested whether similar findings would be found even under a different cultural background. Specifically, in Experiment 2, a similar approach was used in an Asian sample, and there was a neutral reference (a color square) in the neutral condition.

Experiment 2

Self-representation is impacted by cultural context. For instance, westerners normally value the self as independent and less connected to others while East Asians are more likely to value the relationship between the self and others (Markus & Kitayama, 1991). Zhu, Zhang, Fan, and Han (2007) found that processing information in relation to one's mother activated similar neural correlates as self-referenced processing in Asian participants but not in Western subjects, suggesting culture shapes neural correlates of self-representational cognitive processes. Other research found that Eastern Asians tend to focus more on the relationships and contexts when processing pictures such as a lion standing in a desert background while Westerners focus more on the objects such as the lion (Goh, Chee, Tan, Venkatraman, …, & Park, 2007; Nisbett & Masuda, 2003). East Asian participants were tested in Experiment 2 to explore whether there was any cultural difference in self-referenced true and false memories in the DRM paradigm. Since East Asians pay more attention to relationships with others, processing information in relation to others may already show a high level of relational processing. We predicted that East Asians might not exhibit a self-reference false memory effect as strong as Europeans and thus net accuracy (i.e., true memory to true memory plus false memory) should be larger in East Asians.

Method

Participants

Twenty-nine participants from Fudan University, China were tested[13]. The East Asian sample consisted of 13 males and 16 females (M_{age}= 21.2, *SD*= 2.06, min 19 – max 26 years old). All Chinese participants were native Mandarin speakers.

Materials

Fifteen Chinese DRM lists were translated and adapted from Stadler, Roediger, and McDermott (1999). The BAS of each list was measured with a Chinese population in a pilot study based on the procedures in Roediger, Watson, and Gallo (2001). Participants

[13] Post hoc power analysis indicated that the power of Exp. 2 was 0.87 with sample size n = 29.

were first given the Chinese words and were asked to write down the first word it brought to mind for each word. Then the connection strength between each study word (e.g., *garage, drive, road, Mercedes*) and its corresponding critical lure (e.g., *car*) was measured as the probability it elicited the critical lure. BAS was the average probability that the studied items elicited the critical lure. Hence some DRM lists used in the East Asian sample were different from that of the European sample. Fifteen lists were pseudo-randomly assigned to the self-reference (5 lists; *Mean* BAS = 0.19, *SD* =.04), other-reference (5 lists; *Mean* BAS =0.20, *SD* =.04) and the neutral condition (5 lists; *Mean* BAS =0.20, *SD* =.03), with BAS matched in the three conditions, $F(2, 12) = 0.09$, $p = .91$. The recognition list contained 15 critical lures (one per list), 45 studied items (three items per list) and 20 unrelated items. The studied items were from the 1st, 6th, and 10th position of each list.

Design and Procedure

The experiment used a 3 (Reference: self vs. other vs. neutral) × 2 (Memory type: true vs. false) within-subject design. The procedure of Experiment 2 was identical as in Experiment 1 except for the following differences. First, in the East Asian sample of Experiment 2, the name “Li Ming” (a frequently used name in Chinese text books) was used in the “other” condition. Second, in the neutral condition of Experiment 2, a red square appeared together with DRM words. Thus the DRM words appeared together with either their own name or the name “Li Ming” or a red square in the study phase. Other steps in the procedure were exactly the same as in Experiment 1.

Results and Discussion

Recognition Rates

We conducted a 3 (Reference: self vs. other vs. neutral) × 2 (Memory type: true vs. false) repeated measures ANOVA on true and false recognition rates. The results were similar to those in Experiment 1. The main effect of Reference was statistically significant, $F(2, 56) = 25.45$, $p < .001$, partial $\eta^2 = 0.48$. Bonferroni comparisons showed that self-reference led to higher true and false recognition rates than both the other-reference ($p < .001$) and neutral conditions ($p < .001$). The main effect of Memory type was not significant, F (2, 56) = 0.93, p = .34. No significant interaction between Reference and Memory type was found either, F (2, 56) = 1.88, p = .16.

As Figure 7.3b shows, the self-reference false memory effect was found in Experiment 2 as well. Specifically, the self-reference condition had significantly higher false recognition rates than the other-reference condition ($t(28) = 3.19$, $p = .004$, Cohen's d = 0.61) and the neutral condition ($t(28) = 2.45$, $p = .02$, Cohen's d = 0.46). The other-reference condition did not differ from the neutral condition in false recognition rates, $t(28) = 0.96$, $p = .35$. A typical SRE of true memories was also found. A paired-samples t-test showed significant higher true memory rates in self-reference condition than in the other-reference condition, $t(28) = 7.23$, $p < .001$, Cohen's d = 1.46.

Net Accuracy

Net accuracy was calculated for the self-reference, other-reference and neutral conditions. Repeated measures ANOVA showed that there was no significant effect of Reference on net accuracy, $F(2, 56) = 2.12$, $p = .13$. Self-referencing (M = 0.51, 95%CI [0.49, 0.53]) resulted in similar memory net accuracy as in the other-referencing (M = 0.48, 95%CI [0.45, 0.50]) and neutral (M = 0.50, 95%CI [0.47, 0.53]) conditions. In an East Asian sample, it was also found that self-referencing did not improve the overall memory accuracy with DRM paradigm.

Recollection and Familiarity of Self-referenced False Memories

Similar to Experiment 1, we found that Reference impacted recollection scores of critical lures significantly, $F(2, 56) = 14.92$, $p < .001$, partial $\eta^2 = 0.35$, while it did not impact familiarity of critical lures, $F(2, 56) = 0.72$, $p = .49$. Bonferroni comparisons showed that in the self-reference condition, critical lures received significantly higher recollection scores than in the other-reference condition ($t(28) = 4.85$, $p < .001$, Cohen's d = 0.9) and the neutral condition ($t(28) = 4.16$, $p < .001$, Cohen's d = 0.78).

Recollection and Familiarity of Self-referenced True Memories

The East Asian sample had similar results to the European sample inasmuch as in the self-reference condition, presented items had higher recollection scores than in other-referenced condition ($t(28) = 8.17$, $p < .001$, Cohen's d = 1.56) as well as the neutral condition ($t(28) = 6.87$, $p < .001$, Cohen's d = 1.35), and the neutral condition also had higher recollection scores than the other-referenced condition ($t(28) = 3.10$, $p = .004$, Cohen's $d = 0.58$). No significant difference on familiarity was found among the different Reference conditions, $F(2, 56) = 0.44$, $p = .65$ (see Table 7.2).

In summary, Experiment 2 found comparable results in an East Asian sample as with a European sample in Experiment 1. Besides the difference in sample population, Experiment 2 used different DRM lists and a different other-referenced name. However, we still found that self-referencing increased both true and false memories than other-referencing, and the other-referencing condition did not differ from the neutral condition in terms of recognition rates. We also replicated the results that self-referencing did not impact net accuracy of memory, and self-referencing impacted recollection but not familiarity of critical lures and studied items.

Joint Analyses on False Memory

To explore whether the magnitude of the SR false memory effect differed statistically in European and East Asian samples, we combined data of the two samples and treated the Sample as a between-subjects variable. A 2 (Sample: European vs. East Asian) × 2 (Reference: self vs. other) repeated measures ANOVA was conducted on false recognition rates, where Sample was between-subjects and Reference was within-subject. No statistical interaction between Sample and Reference was found, $F(1, 66) = 0.07$, $p = .79$, suggesting Sample did not interact with the SR false memory effect. There was a main effect of Reference, $F(2, 66) = 23.50$, $p < .001$, partial $\eta^2 = 0.26$. That is, in both the European and East Asian samples, self-referencing ($M = 0.78$, 95%CI [0.73, 0.83]) led to higher false memory rates than other-referencing ($M = 0.66$, 95%CI [0.62, 0.71]; $p < .001$, Cohen's $d = 0.59$). There was also a main effect of Sample. East Asian participants ($M = 0.79$, 95%CI [0.73, 0.86]) had statistically more unbiased false memories than European participants ($M = 0.65$, 95%CI [0.60, 0.71]), $F(1, 66) = 11.57$, $p = .001$, partial $\eta^2 = 0.15$.

Joint Analyses on Net Accuracy

A 2 (Sample: European vs. East Asian) × 2 (Reference: self vs. other) repeated measures ANOVA on net accuracy was conducted. No statistical interaction effect between Sample and Reference was found, $F(1, 66) = 0.75$, $p = .39$. There was no main effect of Sample, $F(1, 66) = 0.59$, $p = .44$, indicating no net accuracy differences in European ($M =$

0.51, 95%CI [0.48, 0.53]) and East Asian samples (*M* = 0.49, 95%CI [0.46, 0.52]). No main effect of Reference was found either, F (1, 66) = 0.99, p = .32, suggesting self-referencing (*M* = 0.51, 95%CI [0.49, 0.53]) did not impact net accuracy compared to other-referencing (*M* = 0.49, 95%CI [0.46, 0.52]) across our two samples. Although East Asian participants had different self-referenced true and false memory rates relative to the European sample, self-referencing had a very similar effect on net accuracy in the two samples.

Contrary to our prediction, there was no cultural difference in self-referenced false memories and net accuracies. To note, there are some differences in the experimental materials between the samples, so one needs to be cautious when interpreting the joint analyses. One difference is that the Chinese DRM lists were different from lists used in the European sample. The names used in the other-reference conditions of the two samples were also different, which was due to cultural difference but maybe in future research, the names can be more strictly comparable. Another difference is the number of items in the recognition test. Although the proportions of critical lure in the recognition test were comparable (20% vs. 19% in European and East Asian samples), the different number of studied items per list might impact the results. Nevertheless, the SR false memory effect is based on within-subject comparisons between self- and other-reference conditions and, hence, these small methodological differences are unlikely to have impacted the main outcomes of the current study.

General Discussion

The current study examined whether processing information in relation to oneself (i.e., self-referencing) increased false memories compared to processing information in relation to others (i.e., other-referencing). In two experiments with Eastern and Western samples, we found consistent results that self-referencing enhanced both false recognition rates of critical lures and true recognition rates of studied items in the DRM paradigm. Specifically, self-referencing impacted recollection but not familiarity of studied items and critical lures. In addition, we found that self-referencing did not increase the net accuracy compared to other-referencing and neutral conditions. The current findings support the theoretical view that false memories are not malfunctions of cognitive processes (e.g., Howe, 2011; Schacter, 2012), but that true and false memories may originate from a similar memory mechanism such as spreading activation.

Indeed, the current findings can be readily explained by spreading activation theories of memory (Howe et al., 2009; Roediger et al., 2001), in which relational processing or binding plays a key role. In this theoretical view, nodes (e.g., words, images) and associative links (i.e., relations or bindings) are the cognitive units of memory and they make up the whole memory network (Anderson, 1983; Howe et al., 2009). When participants process the DRM words in relation to themselves, DRM items that share the same theme (i.e., critical lure) are more strongly linked to each other than when DRM words are processed in relation to others. Consequentially, memories for DRM words are better retained under self-referencing and theoretically, activation levels of DRM items in an associative memory network is increased by self-referencing since activation level determines memory rates (Anderson, 1983). Thus, more activation can be spread to related but non-presented critical lures in a memory network, and hence memory rates for critical lures are higher in the self-reference condition. In short, self-referencing increases binding among concepts in a memory network, which leads to more activation spreading along the network to critical lures and this in turn increases false memory rates.

Based on the activation theories, we propose that cognitive processes/factors that increase mnemonic efficiency (by increasing activation level of memory nodes or relational binding among the nodes) may increase susceptibility to associative false memories. The following evidence supports the above proposition. First, our results here have already demonstrated that self-referencing as a superior mnemonic processing leads to high level of false memories. Second, research has found that deeper processing of the information (e.g., defining an item), which is known to enhance true memories, produces more false memories than shallow processing (e.g., counting the letters of a word) (Rhodes & Anastasi, 2000; Thapar & McDermott, 2001). Third, survival processing (e.g., rating items' relevance for a survival scenario) has been found to boost true memories (Nairne & Pandeirada, 2008; Nairne, Thompson, & Pandeirada, 2007), however it has also been found to increase susceptibility to memory illusions (Howe & Derbish, 2010; Otgaar & Smeets, 2010). Fourth, individuals with highly superior autobiographical memories, who can accurately remember everyday details since mid-childhood, are found to be more susceptible to suggestive false memories than normal adults (Patihis, Frenda, LePort, Petersen, …, & Loftus, 2013).

Other evidence has come from the development of relational binding and development of false memories. The ability of binding among stimuli develops with age. Research has found that people's ability to bind isolated parts of pictures improves from childhood to adulthood (Sluzenski, Newcombe, & Kovacs, 2006). Lee, Wendelken, Bunge, and Ghetti (2016) have found that item-item and item-time bindings improve gradually from 8 years of age to adulthood. Interestingly, there is a similar developmental pattern (i.e., the development reversal) of false memory formation from children to adults. That is, children exhibit fewer false memories generated from associative lists such as DRM lists than adults (e.g., Brainerd et al., 2008; Howe et al., 2009; Otgaar, Howe, Brackmann, & Smeets, 2016). More intriguingly, older adults exhibit the phenomenon of hyper-binding in which they make too many associations (Campbell, Hasher, & Thomas, 2010), creating more associative false memories than young adults (Schacter, Koutstaal, & Norman, 1997; Devitt & Schacter, 2016). Taken together, from a developmental perspective, binding ability increases with age from childhood until later adulthood, and susceptibility to associative false memories increases correspondingly.

Relational binding is obviously significant for human memory since it allows individuals to bind arbitrary features of an experience into an integrated episodic memory (Eichenbaum & Cohen, 2004), so are self-referencing, deep processing, and survival processing functional for human memory. However, those cognitive functions inevitably lead to unwanted false memories under certain circumstances. Thus, false memories reflect the adaptive nature of memory in the sense that they originate from efficient functioning of the memory system instead of being malfunctions of the memory system (see also Howe, 2011; Howe & Derbish, 2010; Schacter, 2012).

The self-reference false memory effect found in the current study may not be easily explained by false memory theories such as Fuzzy-trace Theory (FTT, Brainerd & Reyna, 2002; Brainerd, Reyna & Ceci, 2008). FTT proposes that verbatim traces, which refer to the surface features/details of the studied items, lower false memory rates, while more abstract gist traces extracted from studied items increase false memory rates. We and others (e.g., Leshikar et al., 2015) have found that self-referencing increases recollection (i.e., remembered details) of studied items, which corresponds approximately to verbatim traces of studied items. Hence, FTT should predict lower false memory rates under self-referencing since verbatim traces hinder false memories. However, we still found high false memory rates in the self-reference condition. Thus, spreading activation accounts seem to provide a more reasonable explanation for the self-reference false memory effect.

Contrary to our prediction, we did not find any cultural difference in the SR false memory effect and net accuracy. This suggests that the SRE on true and false memories is not sensitive to cultures and may represent a universal memory mechanism. One possible explanation is that spreading activation in a memory network is automatic (Howe, 2005, 2006; Howe et al., 2009; Roediger et al., 2001) and thus not impacted by cultural backgrounds. The automaticity of memory activations implies that activations of critical lures cannot be controlled by consciousness, such as consciously making references to others. Another reason why we did not find a cultural difference may be due to the identity of others used in the other-reference condition. Both "Trump" and "Li Ming" were familiar but not intimate others for our participants. Zhu et al. (2007) have found that when East Asian make references to their mothers, a similar neural pattern as SRE was found. If we use participants' mothers in the other-reference condition in European and East Asian samples, we may find a cultural difference on SR false memories. Further research is needed regarding cultural difference on SRE.

In conclusion, we found a self-reference false memory effect that processing information in relation to oneself led to more false memories than processing information in relation to others (e.g., Trump). The SR false memory effect was stable across western and eastern population samples, with identical effect sizes Cohen's *d* around 0.6. What's more, self-referencing did not increase memory net accuracy across two samples. Based on spreading activation theories of memory, we propose that cognitive processes/factors that increase mnemonic efficiency may increase susceptibility to associative false memories, which reflects the adaptive nature of memory.

References

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of verbal learning and verbal behavior, 22*, 261-295.

Arndt, J., & Reder, L. M. (2003). The effect of distinctive visual information on false recognition. *Journal of Memory and Language, 48*, 1-15.

Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science, 11*, 164-169.

Brainerd, C. J., Reyna, V. F., & Ceci, S. J. (2008). Developmental reversals in false memory: A review of data and theory. *Psychological bulletin*, *134*, 343-382.

Campbell, K. L., Hasher, L., & Thomas, R. C. (2010). Hyper-binding: A unique age effect. *Psychological Science, 21*, 399-405.

Cann, D. R., McRae, K., & Katz, A. N. (2011). False recall in the Deese–Roediger–McDermott paradigm: The roles of gist and associative strength. *The Quarterly Journal of Experimental Psychology, 64*, 1515-1542.

Conway, M. A. (2005). Memory and the self. *Journal of Memory and Language*, *53*, 594-628.

Cunningham, S. J., Brebner, J. L., Quinn, F., & Turk, D. J. (2013). The self-reference effect on memory in early childhood. *Child Development, 85*, 808-823.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*, 17–22.

Devitt, A. L., & Schacter, D. L. (2016). False memories with age: neural and cognitive underpinnings. *Neuropsychologia*, *91*, 346-359.

Eichenbaum, H., & Cohen, N. J. (2004). From conditioning to conscious recollection: memory systems of the brain. New York: Oxford University Press.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

Gallo, D. A., & Roediger, H. L. (2002). Variability among word lists in eliciting memory illusions: Evidence for associative activation and monitoring. *Journal of Memory and Language*, *47*, 469–497.

Goh, J. O., Chee, M. W., Tan, J. C., Venkatraman, V., Hebrank, A., Leshikar, E. D., ... & Park, D. C. (2007). Age and culture modulate object processing and object—scene binding in the ventral visual area. *Cognitive, Affective, & Behavioral Neuroscience, 7*, 44-52.

Gutchess, A. H., Kensinger, E. A., Yoon, C., & Schacter, D. L. (2007). Ageing and the self-reference effect in memory. *Memory, 15*, 822-837.

Hamami, A., Serbun, S. J., & Gutchess, A. H. (2011). Self-referencing enhances memory specificity with age. *Psychology and Aging, 26*, 636-646.

Howe, M. L. (2005). Children (but not adults) can inhibit false memories. *Psychological Science, 16*, 927–931.

Howe, M. L. (2006). Developmentally invariant dissociations in children’s true and false memories: Not all relatedness is created equal. *Child Development, 77*, 1112–1123.

Howe, M. L. (2011). The adaptive nature of memory and its illusions. *Current Directions in Psychological Science, 20*, 312-315.

Howe, M. L., & Derbish, M. H. (2010). On the susceptibility of adaptive memory to false memory illusions. *Cognition, 115*, 252-267.

Howe, M. L., Wimmer, M. C., Gagnon, N. & Plumpton, S. (2009). An associative activation theory of children's and adults' memory illusions. *Journal of Memory and Language*, *60*, 229-251.

Hunt, R. R., & Einstein, G. O. (1981). Relational and item-specific information in memory. *Journal of Verbal Learning and Verbal Behavior*, *20*, 497-514.

Israel, L., & Schacter, D. L. (1997). Pictorial encoding reduces false recognition of semantic associates. *Psychonomic Bulletin & Review, 4*, 577-581.

Klein, S. B. (2012). Self, memory, and the self-Reference effect: An examination of conceptual and methodological issues. *Personality and Social Psychology Review, 16*, 283–300.

Klein, S. B., & Loftus, J. (1988). The nature of self-referent encoding: The contributions of elaborative and organizational processes. *Journal of Personality and Social Psychology, 55*, 5–11.

Kuiper, N. A., & Rogers, T. B. (1979). Encoding of personal information: Self–other differences. *Journal of Personality and Social Psychology, 37*, 499-514.

Lee, J. K., Wendelken, C., Bunge, S. A., & Ghetti, S. (2016). A time and place for everything: Developmental differences in the building blocks of episodic memory. *Child Development, 87*, 194-210.

Leshikar, E. D., & Duarte, A. (2014). Medial prefrontal cortex supports source memory for self-referenced materials in young and older adults. *Cognitive, Affective, & Behavioral Neuroscience, 14*, 236-252.

Leshikar, E. D., Dulas, M. R., & Duarte, A. (2015). Self-referencing enhances recollection in both young and older adults. *Aging, Neuropsychology, and Cognition, 22*, 388-412.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98*, 224.

Mccabe, D. P., Presmanes, A. G., Robertson, C. L., & Smith, A. D. (2004). Item-specific processing reduces false memories. *Psychonomic Bulletin & Review, 11*, 1074-1079.

Nairne, J. S., & Pandeirada, J. N. (2008). Adaptive memory: Remembering with a stone-age brain. *Current Directions in Psychological Science, 17*, 239-243.

Nairne, J. S., Thompson, S. R., & Pandeirada, J. N. S. (2007). Adaptive memory: Survival processing enhances retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 263–273.

Nisbett, R. E., & Masuda, T. (2003). Culture and point of view. *Proceedings of the National Academy of Sciences, 100*, 11163-11170.

Otgaar, H., Howe, M. L., Brackmann, N., & Smeets, T. (2016). The malleability of developmental trends in neutral and negative memory illusions. *Journal of Experimental Psychology: General, 145*, 31-55.

Otgaar, H., Muris, P., Howe, M. L., & Merckelbach, H. (2017). What Drives False Memories in Psychopathology? A Case for Associative Activation. *Clinical Psychological Science, 5*, 1048-1069.

Otgaar, H., & Smeets, T. (2010). Adaptive memory: Survival processing increases both true and false memory in adults and children. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 1010-1016.

Patihis, L., Frenda, S. J., LePort, A. K., Petersen, N., Nichols, R. M., Stark, C. E., ... & Loftus, E. F. (2013). False memories in highly superior autobiographical memory individuals. *Proceedings of the National Academy of Sciences, 110*, 20947-20952.

Rhodes, M. G., & Anastasi, J. S. (2000). The effects of a levels-of-processing manipulation on false recall. *Psychonomic Bulletin & Review, 7*, 158–162.

Roediger, H. L., Balota, D., & Watson, J. (2001). Spreading activation and arousal of false memories. In H. Roediger, J. Nairne, & A. Surprenant (Eds.), *The nature of* *remembering: Essays in honor of Robert G. Crowder. Science conference series* (pp. 95–115). Washington, DC: American Psychological Association.

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*, 803–14.

Roediger, H. L., Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review, 8*, 385-407.

Rogers, T. B., Kuiper, N. A., & Kirker, W. S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology, 35,* 677–688.

Rogers, T. B., Rogers, P. J., & Kuiper, N. A. (1979). Evidence for the self as a cognitive prototype: the "false alarms effect". *Personality and Social Psychology Bulletin, 5*, 53-56.

Rosa, N. M., Deason, R. G., Budson, A. E., & Gutchess, A. H. (2015). Self-referencing and false memory in mild cognitive impairment due to Alzheimer's disease. *Neuropsychology*, *29*, 799-805.

Rosa, N. M., & Gutchess, A. H. (2013). False memory in aging resulting from self-referential processing. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 68, 882-892.

Schacter, D. L. (2012). Adaptive constructive processes and the future of memory. *American Psychologist*, *67*, 603-613.

Schacter, D. L., Koutstaal, W., & Norman, K. A. (1997). False memories and aging. *Trends in cognitive sciences*, *1*, 229-236.

Serbun, S. J., Shih, J. Y., & Gutchess, A. H. (2011). Memory for details with self-referencing. *Memory*, *19*, 1004-1014.

Sluzenski, J., Newcombe, N. S., & Kovacs, S. L. (2006). Binding, relational memory, and recall of naturalistic events: A developmental perspective. *Journal of Experimental Psychology: Learning Memory and Cognition, 32,* 89-100.

Stadler, M. A., Roediger, H. L., & McDermott, K. B. (1999). Norms for word lists that create false memories. *Memory & cognition*, *27*, 494-500.

Sui, J., & Humphreys, G. W. (2015). The integrative self: how self-reference integrates perception and memory. *Trends in Cognitive Sciences*, *19*, 719-728.

Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: A meta-analysis. *Psychological Bulletin, 121*, 371–394.

Thapar, A. & McDermott, K. B. (2001). False recall and false recognition induced by presentation of associated words: effects of retention interval and level of processing. *Memory & Cognition, 29*, 424-432.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46*, 441–517.

Yonelinas, A. P., & Jacoby, L. L. (1995). The relation between remembering and knowing as bases for recognition: Effects of size congruency. *Journal of Memory and Language, 34*, 622–643.

Zhu, Y., Zhang, L., Fan, J., & Han, S. (2007). Neural basis of cultural influence on self-representation. *Neuroimage, 34*, 1310-1316.

Appendix 7A: DRM lists used in Experiment 1

Sleep	Music	Bread	Smoke	City	Soft	Long	Fruit	Needle
bed	sound	butter	cigarette	town	hard	short	apple	thread
rest	piano	food	puff	crowded	light	narrow	vegetable	pin
awake	sing	eat	blaze	state	pillow	time	orange	eye
tired	radio	sandwich	pollution	capital	plush	far	kiwi	sewing
dream	band	rye	ashes	streets	cotton	hair	citrus	sharp
wake	melody	jam	cigar	subway	fur	island	ripe	point
snooze	horn	milk	chimney	country	touch	road	pear	prick
blanket	concert	flour	fire	New York	fluffy	thin	banana	thimble
doze	instrument	dough	tobacco	village	feather	distance	berry	haystack
snore	symphony	slice	pipe	metropolis	downy	line	cherry	thorn
nap	orchestra	loaf	lungs	suburb	kitten	low	basket	injection
drowsy	rhythm	toast	flames	urban	tender	rope	cocktail	syringe

Sweet	Foot	Pen	Car	Chair	Window	Shirt	Black	River
sour	shoe	pencil	truck	table	door	blouse	white	water
candy	hand	write	bus	sit	glass	sleeves	Dark	stream
sugar	toe	fountain	automobile	seat	pane	pants	charred	lake
bitter	kick	leak	vehicle	couch	shade	tie	night	Mississipppi
taste	sandals	quill	drive	desk	ledge	button	funeral	tide
nice	walk	Bic	jeep	recliner	sill	shorts	colour	swim
honey	ankle	scribble	Ford	sofa	house	polo	grief	flow
soda	arm	crayon	keys	cushion	curtain	collar	death	barge
chocolate	boot	cross	garage	swivel	view	pocket	ink	creek
cake	inch	marker	highway	stool	breeze	belt	coal	brook
tart	sock	cap	van	rocking	screen	linen	brown	fish
pie	knee	letter	taxi	bench	shutter	cuffs	grey	bridge

汽车	国旗	女孩	生气	小偷	寒冷	针尖	山脉	面包
车库	飘扬	男孩	激怒	警察	冬天	丝线	起伏	黄油
驾驶	五星	漂亮	愤慨	偷盗	寒战	大头针	陡峭	吐司
公路	旗杆	裙子	大怒	偷窃	寒风	针眼	攀登	面粉
奔驰	升起	细腻	狂怒	盗贼	颤抖	缝纫	险峻	果酱
轿车	国歌	美丽	怒火	贼	结冰	锐利	顶峰	早餐
修车厂	旗帜	可爱	愤怒	窃贼	冰	尖锐	小山	薄片
出租车	红色	约会	暴怒	强盗	霜冻	扎破	丘陵	生面团
吉普车	升旗	舞蹈	情绪	夜贼	气候	顶针	登山者	三明治
货车	象征	长发	憎恨	骗子	下雪	荆棘	山岗	奶油
车辆	标语	辫子	讨厌	抢夺	霜	刺痛	高峰	牛奶
车厢	敬礼	洋娃娃	脾气	抢掠	天气	注射	雄伟	热狗
驾驶员	中国	伶俐	争论	罪恶	暖和	编制	山顶	食物

男人	甜的	皇帝	医生	垃圾	钢笔
女人	糖果	康熙	诊所	废物	笔帽
英俊	糖水	臣民	听诊器	扔弃	墨水
潇洒	白糖	皇后	护士	废品	笔尖
健壮	食糖	王位	内科	堆积	文具
胡子	奶糖	加冕	门诊	废料	水笔
肌肉	蜂蜜	皇宫	外科	废弃	笔袋
强壮	蜜饯	君王	医师	垃圾站	铅笔
胡须	冰糖	宝座	临床	清扫	勾划
丈夫	苦的	统治	处方	破烂	纸张
男性	酸的	王冠	诊断	废话	书写
西装	巧克力	君主	治疗	破旧	鹅毛
雄性	点心	宫殿	患者	清洁工	毛笔

CHAPTER 8

General Discussion

A SUMMARY OF THE FINDINGS

The general aim of the current dissertation was to examine the consequences of belief and recollection on different cognitive-behavioral tasks. I was particularly interested in the question regarding which component of memory, belief or recollection, was more critical in impacting cognitive-behavioral performance. In **Part I**, I first reviewed the negative consequences of false memories in the Chinese legal context (Chapter 2) and examined the positive consequences of true and false memories on a cognitive task (Chapter 3). Specifically, Chapter 2 described the devastating consequences of eyewitness false memories in the court room in that such faulty testimonies could lead to the imprisonment of innocent people. Eyewitness false memories, either originating from external misinformation or internal mechanisms (e.g., spreading activation), are deeply rooted in the malleable nature of memory and understanding the constructs of true and false memory is imperative. Chapter 3 studied the nature and cognitive-behavioral consequences of true and false memories for pictures. It was found that both true and false memories can facilitate performance on a perceptual closure task, and three independent experiments convincingly showed that false memories for visual stimuli contain conceptual but not perceptual mental representations.

After demonstrating both negative and positive consequences of memory, **Part II** examined the role of belief and recollection, which are the primary two constructs of memory (James, 1890; Rubin, 2005; Scoboria et al., 2014) in three cognitive-behavioral paradigms: a problem solving paradigm, a decision making paradigm, and a food preference paradigm. Chapter 4 investigated whether undermining belief in true and false memories would impact related problem solving performance. In two experiments, it was demonstrated that when belief was undermined, problem solving rates also dropped and this effect was even evident one week later. In a decision making paradigm (Chapter 5), it was found that challenging genuine memories via false feedback decreased both recollection and belief ratings for memory associations, and subsequently led to the disappearance of memory related decision biases. Chapter 6 examined which component of memory is accountable for people's food preferences. Results from Chapter 6 showed that belief, but not recollection regarding a food aversive event, determined people's preferences for that particular food. Moreover, explicit self-reported belief, and not implicit belief, impacted food preferences. Taken together, these studies provide converging and crucial evidence that belief is more important than recollection in impacting people's behavior.

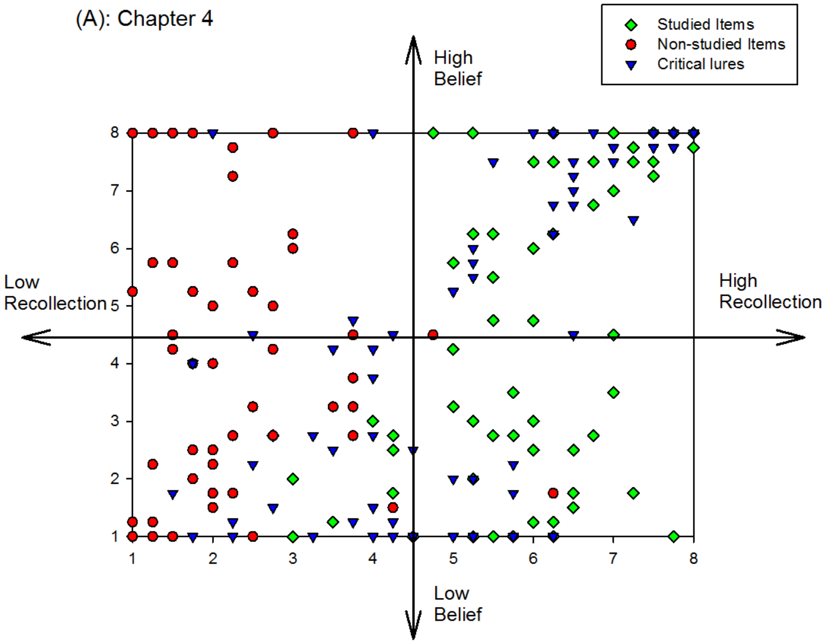
One unique feature of episodic memory is that it is *autonoetic* (Tulving, 2002). Autonoesis refers the ability to place ourselves in the past to accomplish the process of remembering, and without the self, remembering is not possible. Chapter 7 investigated memory mechanisms with respect to the self, i.e., the self-reference effect. Chapter 7 found that processing information related to oneself (self-referencing) increased true memories as well as false memories compared to processing information related to others (other-referencing). Chapter 7 supported the theoretical position that the self facilitates relational processing or the binding among different stimuli, which exhibits the constructive or malleable nature of memory. This binding not only facilitates the processing of true memories, but also increases the likelihood to false memory formation.

I will now discuss how the above-mentioned findings support belief and recollection as dissociated constructs of memory. Then I will discuss the respective cognitive-behavioral consequences of belief and recollection, and discuss potential mechanisms that underlie the observed findings. Finally, implications of the research described in this dissertation will be discussed.

CONSTRUCTS OF MEMORY

Dissociation of Belief and Recollection in Memories for Stimuli

In Chapters 4, 5, and 6, belief and recollection for different experimental stimuli (words, pictures, and autobiographical events) and different memory types (true memory and false memory) were studied. Dissociation between belief and recollection was consistently found, supporting the theoretical view that these two are dissociated constructs of memory (Rubin, 2005, 2006; Scoboria et al., 2014). The dissociated model of belief and recollection proposes that mental representations of events can represent different levels of belief and recollection dimensions, leading to different event categories such as believed memories, nonbelieved memories, believed but not remembered events, and nonbelieved, not remembered events. As Figure 8.1 shows, the different experimental manipulations in Chapters 4 and 6 successfully elicited all possible categories of memories based on recollection and belief ratings, confirming the dissociated model of memory constructs.



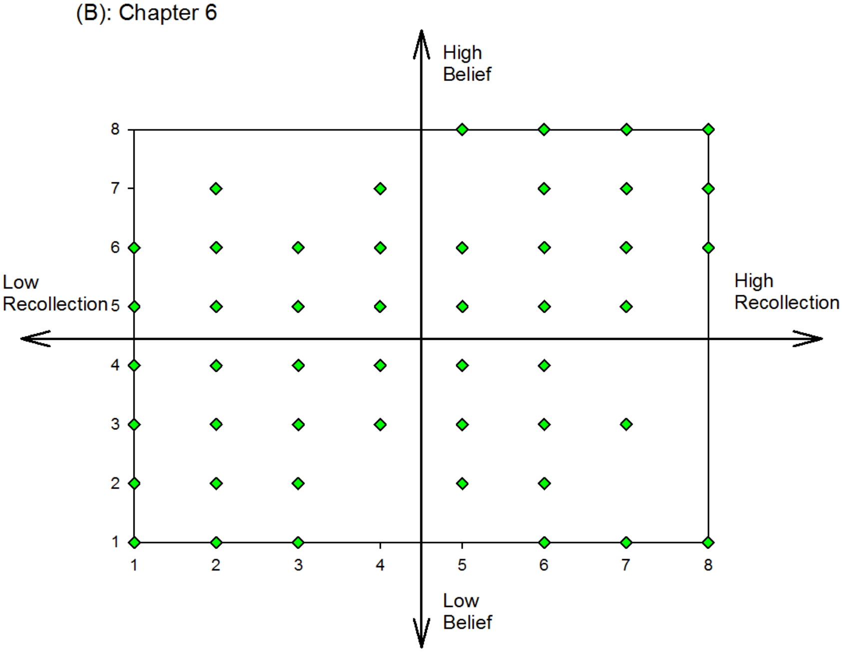


Figure 8.1. (A) Participants' (N = 34) average recollection and belief ratings (1-8) for multiple words including studied items, critical lures and unrelated non-studied items. (B) Participants' (N = 161) recollection and belief ratings (1- 8) for one (false) autobiographical event, i.e., being sick after eating egg salad as a child. Note that one dot in the figure may represent multiple participants with the same rating.

From a dichotomous perspective, the manipulation described inChapter 6 also successfully elicited all possible combinations of belief and recollection of an autobiographical event (being sick after eating egg salad in childhood). After a false suggestion and partial debriefing, participants reported whether they had a recollection (yes vs. no) and a belief (yes vs. no) in the critical autobiographical event, and 28% of the participants (n = 45) reported to hold believed memories, 13% (n = 21) had nonbelieved memories, 16% (n = 26) had beliefs but no recollections, and 43% (n = 69) had no belief and no recollection of the critical event.

Association of Belief and Recollection in Memory Associations

In Chapters 4 and 6, feedback challenging participants' memories led to a significant decrease of belief whereas recollective aspects of memory remained relatively intact. However, in Chapter 5, after false feedback was provided to undermine participants' memory associations (that two stimuli were not paired together), participants' belief ratings as well as their recollection ratings for the associations decreased significantly. An additional repeated measures ANOVA showed that there was no interaction effect between feedback and memory components, $F(1, 40) = 3.93$, $p > .05$, indicating that false feedback to challenge memories exerted a comparable decrement in belief and recollection ratings. As Figure 8.2 shows, there were few nonbelieved memories (low belief with high recollection) since recollection rating also dropped when belief was undermined via false feedback. This suggests that recollections for memory associations between stimuli were much more malleable than recollections for discrete stimuli.

One possible reason might be that recollections for discrete stimuli and recollections for memory associations are fundamentally different. For recollections of discrete stimuli such as a word or a picture, one can recall sensory details such as the color, size, or emotion it evokes. However, for memory associations such as a circle temporally paired with a square, apart from the recollections for the circle and the square respectively, there is hardly any sensory detail left in a mental representation of the "association". Indeed, from a memory network perspective (Anderson, 1983; Howe et al., 2009), a memory network consists of nodes that refer to mental representations of objects, and also relations or associations among the nodes. Chapter 5 showed that remembering an association and believing an association appear to be entrenched within each other. This suggests that in the case of memory associations, a possible exception is detected for the dissociated model of belief and recollection.

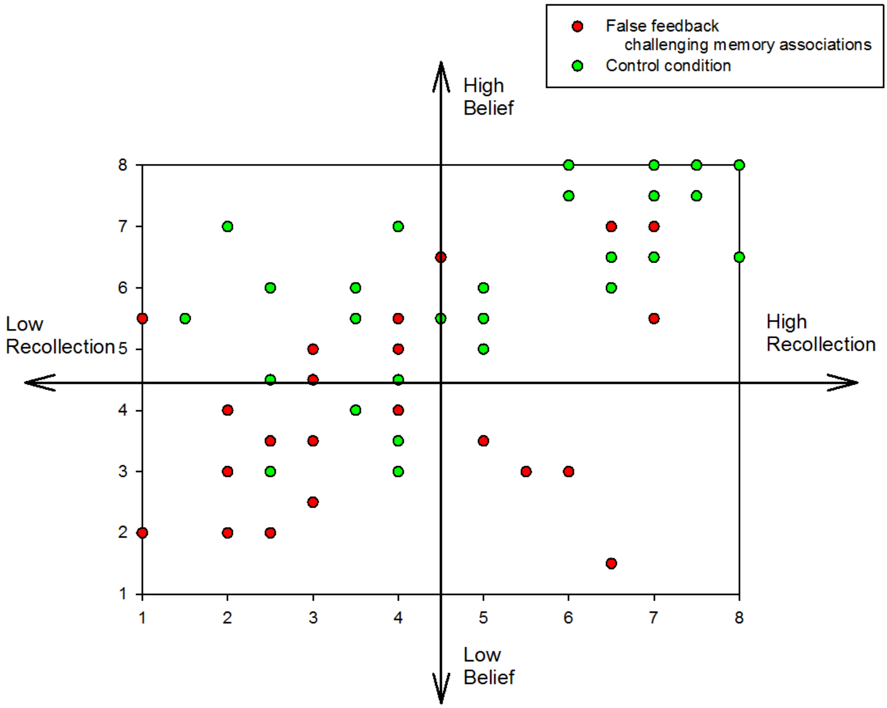


Figure 8.2. Participants' (*N* = 41) recollection and belief ratings (1- 8) distribution for memory associations (data from Chapter 5).

COGNITIVE-BEHAVIORAL CONSEQUENCES OF BELIEF VERSUS RECOLLECTION

Belief Determines Behavior

Across five experiments using three cognitive-behavioral paradigms, it was consistently found that change of belief led to change of behavior. Table 8.1 illustrates the effect sizes of manipulating belief on cognitive-behavioral outcomes in all studies. More specifically, Experiments 1 and 2 in Chapter 4 used an insight-based problem solving task (i.e., the CRAT) where participants were required to solve a three word puzzle (e.g., *board/mail/magic*) by coming up with a word that could link to all three words (i.e., *black*).

Previous research demonstrated that solving CRAT problems was facilitated by the previous presentation of the solutions and false memories of the solutions (Howe et al., 2010; 2013; 2015). In the current work, for both true (Exp. 2) and false memory (Exps. 1 & 2), problem solving rates decreased significantly compared with believed memories when belief was undermined.

Table 8.1

Effect sizes of the belief effect in current experiments

	Cognitive-behavioral measurement	Memory type	Cohen's *d* (*Belief vs. No/undermined Belief*)	Size of the effect*
Chapter 4; Exp. 1	Problem solution rates	False	0.51	Medium
		True	- 0.47ǂ	Medium
Chapter 4; Exp. 2	Problem solution rates	False	1.19	Large
		True	0.96	Large
Chapter 5	Decision making rates	True	0.74	Medium to Large
Chapter 6	Food preference scores	False	0.58 (Egg salad)	Medium
			0.65 (Boiled eggs)	Medium to Large
Weighted mean effect size			0.66	Medium to Large

*The size of the effect was interpreted as small (*d* = 0.2), medium (*d* = 0.5) and large (*d* = 0.8) based on the criteria by Cohen (1988).

ǂ This effect size was negative because the results showed that undermining belief for true memories increased problem solution rates. Thus it still showed that manipulating belief impacted people's behavior, which is consistent with other findings, but only the belief effect worked in a different way.

The experiment in Chapter 5 adopted a classical conditioning decision-making task. Participants first learned associations of a picture and a circle paired together and later the circle was rewarded with money in a separate phase. In previous studies of the consequences of learned associations, memories for the associations would enable participants to choose the picture that used to be paired with the rewarded circle to win money, even though it was never rewarded (i.e., the sensory preconditioning effect; Brogden, 1939; Walther, 2002; Wimmer & Shohamy, 2012). The experiment described in Chapter 5 manipulated participants' memory associations via false feedback telling participants that the picture was *not* paired with a rewarded circle. Both recollection and belief ratings for the associations were attenuated, and participants no longer exhibited decision preferences for the pictures. As discussed previously, the results that recollection ratings decreased through receiving such false feedback was not in line with our expectation. Hence, we were unfortunately unable to identify whether it was recollection or belief for the associations that impacted later decision making. Nonetheless, it showed that feedback manipulation aimed at undermining belief in memory associations can change people's decision making behavior.

The experiments described in Chapter 6 directly compared the consequences of belief and recollection on food preferences. An abundance of research has shown that memories of past food experiences, such as getting sick after eating egg salad, affect food preferences and eating behavior for that particular food (e.g., Berkowitz et al., 2008; Bernstein et al., 2005, 2009; Scoboria et al., 2008, 2012). Chapter 6 made a first attempt at distinguishing the roles of belief and recollection of past food experiences in impacting people's food preferences. By creating groups with different combinations of belief (yes vs. no) and recollection (yes vs. no) status for the critical event (getting sick on egg salad), Chapter 6 demonstrated that belief in the food experience impacted participants' egg salad preferences. Recollection of the negative egg-salad experience, however, did not impact food preferences. Moreover, it was explicit self-reported belief, rather than implicit belief, that changed participants' food preferences.

Taken together, the results from the various studies that were discussed above strongly point to the direction that belief matters more than recollection in impacting cognitive-behavioral performance such as problem solving, decision making, and food preferences. For ease of referring to this finding in the following text, I have termed this effect the *belief effect*. As Table 8.1 shows, the effect sizes of the belief effect ranged from medium to large in the current experiments. The weighted mean Cohen's d of the belief effect was 0.66, implying a rather strong effect of belief in memory related cognition and behavior.

The next question would be: what is the function or consequence(s) of recollection then? For example, one existing view is that recollection of past experience allows us to imagine and plan the future (Schacter & Addis, 2007). Another recent theoretical view proposed that recollection allows people to verify their beliefs about the past and to communicate with others for reasons why they hold certain beliefs (Mahr & Csibra, 2018). Indeed, if there exists recollection, it should have an impact on some related components. Just like the appendix that is found to be useless in the human digestive system, it can still cause severe pain in the abdomen. The relevant question is then what kind of consequences recollection could still exert on our behavior given that behavior is primarily guided by our beliefs.

Recollection Impacts Fluency

One possible way to unravel the precise consequences and functions of recollective experiences is to look more closely at the reaction time and phenomenological data of the experiments described in this dissertation. In Chapter 4, Experiment 1, it was found that for studied items, undermining belief in their occurrence led to faster solution times in solving the CRAT problems than the control condition. It means that recollections for the studied items made participants find the solution more quickly, although recollections in general did not increase the number of CRAT problems that were solved. Otgaar, Moldoveanu, Wang, and Howe (2017) also found that nonbelieved false memories for words (e.g., bread) led to faster identification of the same blurred words (e.g., blurred word “bread”) than believed-not-remembered words.

In Experiment 1a of Chapter 6, recollection was found not to impact individuals’ preferences of egg-containing foods, but recollection did impact reaction times of processing egg salad-related statements. Participants with recollections of getting sick on egg salad categorized recollection related sentences such as “I got sick after eating egg salad” faster than sentences such as “I never got ill after eating egg salad” (Cohen’s $d \sim 0.5$) in the aIAT. When looking more closely at the phenomenological aspects of participants’ recollections, the results showed that every recollection aspect that was measured, including visual details, smell, taste, setting, time, negative emotion and coherent story, was related to reaction times of categorizing related sentences (ps $< .05$; Cohen’s d $0.5 \sim 0.9$). A regression model showed that the average score of phenomenological aspects of recollection explained almost 10% of a reaction time index (D score in aIAT). Thus, the more detailed and vivid participants' recollections of a past experience were, the faster participants were in processing statements consistent with their recollections.

Based on the above findings, it could be argued that recollection facilitates the processing of fluency of recollection-related information. Processing fluency refers to the ease (or difficulty) of processing information, such as the speed to identify the physical form of a stimulus after pre-exposure (Lewandowsky et al., 2012; Schwarz, Sanna, Skurnik, & Yoon, 2007). In episodic memories, different and random concepts (e.g., egg salad, physical symptoms of getting sick) are combined together into a coherent experience (Tulving, 2002), which makes later processing of that mental image or a statement that describes the mental image smooth and fluent. Fazio and co-workers (2013) showed that after people read stories with errors (e.g., "Franklin invented the light bulb"), they would still answer that "Franklin" invented the light bulb even when they knew that this was incorrect. Such findings could be explained by postulating that presenting the sentence "Franklin invented the light bulb" before subjects had to answer certain critical questions ameliorated the fluency of the message. In sum, what I would like to propose here is that recollection – independent from belief – may play an influential role in guiding behavior in a more automatic manner through impacting the fluency of processing. Future research examining the function and consequences of recollection may investigate whether recollection in and of itself is sufficient to increase the fluency of experiences.

Although recollection is speculated to drive fluent processing, the current dissertation's findings suggest that recollection does not seem to contribute much to behavior that requires an individual to consciously solve a problem, make a decision to win money, or rate how much they like a certain type of food (cf. Chapters 4, 5, and 6). The critical issue is why belief rather than recollection is most influential in determining an individual's behavior. The dissociated model of belief and recollection implies to some extent that these two components might have different functions and exhibit different impacts on behavior since belief and recollection are different constructs of memory. However, it does not provide an exhaustive explanation about why belief and recollection have different cognitive-behavioral consequences. In the following section, I will discuss possible mechanisms underlying the proposed belief effect.

POTENTIAL MECHANISMS

To some degree, the finding that belief seems to play a potent role in cognitive-behavioral performance might be viewed as surprising. The reason for this is because traditional views on memory mostly agree that recollection, i.e., mentally re-experiencing an event, is the defining characteristic of memory (e.g., James, 1890; Rubin, 2006; Tulving, 2002). In contrast, the results described in the various chapters of this dissertation clearly show that recollections of past experiences are perhaps not so central to memory as belief is. Here I discuss possible mechanisms involved in the belief effect by combining the dissociated view of recollection and belief with other memory frameworks and learning and decision making theories.

Spreading Activation

Memory can be seen as a network of information comprising of nodes and associations or bindings among the nodes (Anderson, 1983; Howe et al., 2009). Nodes are mental representations of perceived stimuli, which can entail semantic, imagery, phonological information and so on. Episodic memory binds related perceived information into a coherent and meaningful mental representation of a past event (Eichenbaum & Cohen, 2004; Tulving , 2002). For example, as Figure 8.3 shows, a hypothetical memory representation of the event "being sick on egg salad" mainly binds egg salad with being sick symptoms, along with other stimuli representations such as bread, a table, a room, music played on TV, etc. The mental process of remembering is thus the activation of a web of nodes that contribute to the feeling of recollection. For example, when a cue is presented such as a question "how is your holiday in Nice?", the cue activates related memory nodes such as beach and French food. Spreading activation theories of how a memory network works (Anderson, 1983; Howe et al., 2009; Otgaar et al., 2017; Roediger et al., 2001) suggest that activation of a particular node will spread automatically to its nearby associated nodes. For instance, the presentation of the word "egg salad" will activate the "egg salad" node and its activation will spread to its nearby nodes.

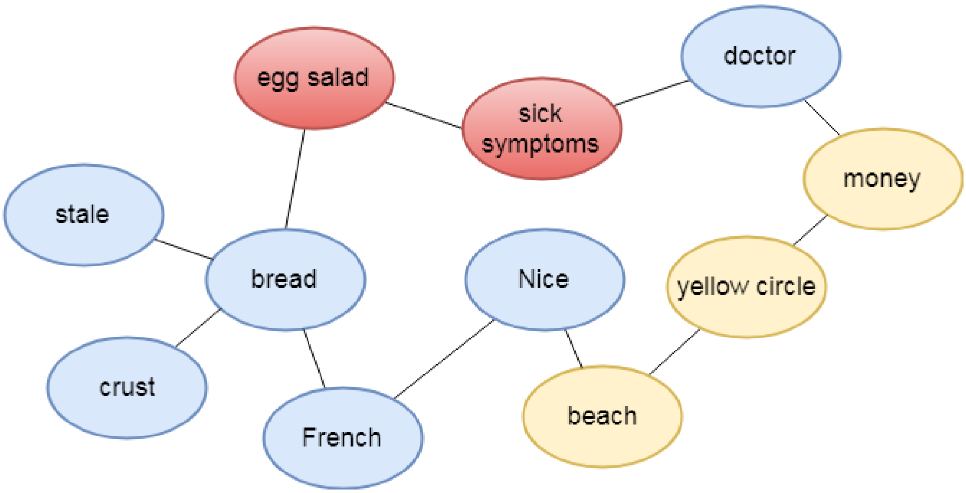


Figure 8.3. Schematic demonstration of a memory network. Circles represent nodes and lines represent associations. Different colors of circles represent different episodes of memory.

In learning theories such as classical conditioning, association is a basis for learning in that it connects two stimuli by (repeated) pairing or co-exposure of the stimuli (e.g., Klein, 2011). As in the Garcia effect, i.e., conditioned taste aversion (Garcia & Koelling, 1966), rats avoided drinking tasty sweet water after sweet water was paired with nauseous toxic LiCl. In our study, when egg salad is suggested to have been paired with sickness such as pain in the stomach, an association between egg salad and being sick symptoms is thus established. Combining the behavioral learning perspective with the memory network theories, the egg salad - sickness association can be encoded in individuals' memory network. Thus, when participants are asked to rate their preferences for egg salad, the word "egg salad" activates the "egg salad" node in one's memory network and the activation then spreads automatically to the "sick symptoms" node, which could make participants have a lower preference for egg salad.

This mechanism can explain the results obtained via other cognitive-behavioral paradigms throughout this dissertation as well. In the priming effect of memories on problem solving, studied words (e.g., the word "bread") are presented and they activate corresponding nodes (i.e., the node "bread"). When a puzzle (e.g., Crust/stale/French) whose solution is "bread" is presented, previous activation of "bread" makes it more accessible and easier to retrieve, enabling participants to have a higher chance to generate the solution word “bread”. In the sensory preconditioning effect studied in Chapter 5, the activation of money could spread to a non-rewarded but associated picture, thus making participants choose those pictures to win money (see similar discussions in Morewedge & Kahneman, 2010).

The next relevant issue is: what is the role of belief in the above mentioned findings? Why does undermining or retracting belief result in an attenuation of the priming effect, the preconditioning effect, and the food preference effect? One possibility that I want to put forward here is that belief functions as a monitoring mechanism in the spreading activation process.

Reality Monitoring

Belief in the occurrence of an event is a metacognitive judgment of past experiences (Scoboria et al., 2014). Such a metacognitive judgment can be viewed as reminiscent of the idea of *reality monitoring* or *source monitoring* within the source monitoring framework (SMF; Johnson, Hashtroudi, & Lindsay, 1993; Mitchell & Johnson, 2009). In the SMF, source monitoring refers to a set of processes involved in making attributions of the origins of memories: for example, distinguishing thoughts, imaginations, etc., from memories of perceived events (Johnson et al., 1993). When people experience an event, mental representations will be recorded as features of event memories including perceptual information (e.g., color and size), spatial information, temporal details, emotional information and so on. Besides perceived events, however, mental representations can also originate from imaginations, dreams, or hallucinations. Reality monitoring attributes certain mental representations in a rapid and non-deliberative way to events that happened or to mere imaginations/dreams (Mitchell & Johnson, 2000). Characteristics of the mental representation (i.e., vividness, coherence of a *recollection*) as well as other factors (e.g., biases, knowledge, etc.) can all possibly affect the final judgement of the authenticity of a mental representation. For example, one might conclude a mental episode does not refer to a memory because “I should have more detailed memories if it indeed happened” (Johnson et al., 1993).

Of relevance and interest here is that the reality monitoring mechanism shares a lot of similarities with the belief judgment of a recollection whether it occurred or not. One thing to note, though, is that belief judgment can also monitor the authenticity of experiences with no recollections, which is different from the source monitoring in SMF. Nonetheless, the process that people retract their beliefs in recollections essentially represents re-attributions of the recollections to sources other than reality. For example, Piaget stopped believing in being kidnapped because he found out that it was originated from his nurse's fabrication. Oliver Sacks no longer believed being bombed because he realized that his memory was originated from a relative's letter. To stop believing in an autobiographical memory implies that the corresponding recollection is re-attributed to unfounded origins. For instance, there are alternative attributions such as dreams, imaginations, and happening to others, or people retract beliefs because the internal features of event representations do not match other memories (Scoboria et al., 2015). For example, in Chapter 1, a case was described in which a person retracted a recovered memory of child sexual abuse because "*The recovered memories were crystal clear in every detail (that's how I realized they were not true) ... my real memories faded with time unless I saw photos or talk with relatives*" (Ost, Costall, & Bull, 2002). This perspective dovetails nicely with a recent proposition by Blank (2017), who argued that belief acts like a monitoring "reality check" to balance recollections and the reality. Thus, the belief judgment could be viewed as a process to make attributions of mental representations, i.e., to decide whether they belong to a memory network or not.

Hence, belief may be involved in the monitoring process of the various associations in an autobiographical memory network. For recollections that did not occur, they are attributed to imaginations, dreams, or events that happened to others rather than to an autobiographical memory episode and those recollections may gradually become detached to other nodes in the memory network. Indeed, research has found that believed memories are rated as having greater personal significance and a greater connectedness to other life events than nonbelieved memories (Mazzoni et al., 2010). Belief thus may have enhanced establishing associations between memory nodes. The study in Chapter 5 also found that belief was closely related to the association between two stimuli. When belief in the association decreased, participants were less likely to remember the association. Similarly, in Chapter 6, the finding that participants did not believe that they had been sick after eating egg salad in childhood might imply that there was no association between egg salad and sick symptoms. Moreover, when participants believed the occurrence of the sick-on-egg salad event, a mental link between egg salad and sickness may have formed. As a result, they rated their preferences for egg salad lower than participants with no beliefs in the egg salad- sickness association.

Memory as a Flexible and Reconstructive System

Chapter 3 and Chapter 7 found that the presence of related stimuli such as DRM word lists or pictures resulted in false memories of related but non-presented stimuli. Those false memories were accompanied with strong recollection (Exps. 1 & 2, Chapter 7; see also McCabe, Roediger, McDaniel, & Balota, 2009; Roediger, Watson, Gallo, & McDermott, 2001). These results can be easily explained by the spreading activation and reality monitoring mechanisms of memory (see Roediger et al., 2001; Johnson et al., 2000). When related words such as *butter*, *food,* and *sandwich* are presented, their corresponding nodes are activated and activation spreads to the "bread" node. Activation of the bread node can evoke a feeling of recollection and when reality monitoring fails to attribute it as non-presented item, participants indicate that the concept bread was presented and produce a false memory that contains similar characteristics as a true memory would have.

Chapter 4 and Chapter 6 showed that the belief in occurrence for true and false memories can easily be altered via social feedback, possibly leading to a change of memory associations as was previously discussed. Correspondingly, Chapter 5 demonstrated that memory associations can be re-formed via social feedback. Taken together, memory associations in an episodic memory network can be shaped and reconstructed by incorporating new information. This idea is in line with research showing that memory can be edited, that is, people sometimes "borrow" or recombine content from true memories to construct vivid false memories (Brown, Caderao, Fields, & Marsh, 2015; Lampinen, Meier, Arnal, & Leding, 2005; Lampinen, Ryals, & Smith, 2008). Belief may play an important role in reconstructing memory associations and it is much more flexible than recollection. The dissociation of autobiographical belief and recollection as well as the flexibility of belief enable people to cope with an environment that is full of changes and novel situations. Think about a scenario that belief is strictly tied with recollection, that is, one only believes in the events that one has experienced just like the old saying "*seeing is believing*" does. In that case, one cannot get much useful experience from memory since people rarely encounter the exact same situation as before. By actively incorporating feedback to memories and information referring to possible memory associations, one can cope with novel situations in an adaptive way.

IMPLICATIONS AND FUTURE DIRECTIONS

Memories can have significant ramifications in the legal arena. The criminal justice system relies heavily on eyewitness testimony, which is assumed to be based on eyewitness' memory. This is especially the case when forensic technical evidence such as DNA samples are lacking. Since 1989, there have been thousands of cases where suspects were prosecuted, until DNA testing proved that they were wrongly accused. Seventy-two percent (n = 242) of these DNA exoneration cases in the US based Innocence Project were victims of mistaken eyewitness identification (http://www.innocenceproject.org/causes-wrongful-conviction). Among legal professionals, it is oftentimes assumed that eyewitnesses testify and make identifications based on their "memories" of what they experienced. There is research showing that knowing judgments of events (i.e., feeling of familiarity but no details recalled) are more likely to predict high accuracy than recollection (Sauerland & Sporer, 2009). However, at present, it is unknown whether eyewitnesses make identifications based solely on their beliefs, their recollections, or both.

There are several reasons to argue that belief in the occurrence of an event also plays an important part in eyewitness testimony. Memory researchers have often failed to distinguish between memory and belief (Scoboria et al., 2004). Research on the consequences of belief and recollection may shed light on how witnesses' reports may be altered under certain circumstances. For instance, information from one witness could be indirectly passed to another witness through a third party, such as a police officer, who informs the witness about what another witness had said (Luus & Wells, 1994). When the information provided by the police officer contradicts the witness's memory, the witness may undermine his belief in his recollection and behave in accordance with the belief effect. In the memory conformity effect (Gabbert, Memon, & Allan, 2003; Wright, Self, & Justice, 2000), witnesses' reports are changed after receiving feedback from co-witnesses. It is unknown whether their change of reports is a change of beliefs or change of memories.

What we need is research on this type of dynamic: how social feedback leads to change of beliefs/recollections and further influences behavioral output in the legal area.

The current results also have the potential to illuminate the links between recollection, belief, and symptoms in certain populations, such as people with a traumatic history (e.g., retractors of abusive memories, PTSD patients). A factor worth noting for why people undermine their beliefs in the occurrence of traumatic experiences is that they have a strong desire to not remember the event. In Scoboria et al.'s (2015) research, a small percentage of college respondents (4%) were uncomfortable with or disliked the content of their memories and tried to "push it away from my mind" or "did not want to believe that that happened" (p. 554). They successfully compelled themselves to withdraw belief, and thus formed nonbelieved memories. Intriguingly, people who have experienced traumatic events frequently encounter involuntary, intrusive memories of those events (Berntsen, 2010) and often adopt the strategy of denying or undermining their belief in the occurrence of traumatic events (Horowitz, 1986). According to the belief effect found here, if undermining belief undermines rates of memory intrusions as well, developing methods to change beliefs would be a meaningful research direction with practical significance.

Interestingly, experimental research has found that people can "repress" unwanted memories by not thinking about the memories. In the Think/No Think paradigm (Anderson, 2003; Anderson & Green, 2001), participants first learned word pairs such as *street-book* in which the second word was the target word. Next, participants were instructed not to think about the target word (i.e., *book*) when the cue word (i.e., *street*) was presented (repression manipulation). Results found that participants exhibited significantly worse memories for the target words in the repression condition than in the baseline condition where word pairs were not presented in the no-think stage. Research also showed that the repression effect was especially stronger for negatively emotional materials than for neutral materials (Lambert, Good, & Kirk, 2010). Related to the phenomenon that people tried to push away uncomfortable memories from mind resulting in nonbelieved memories, one interesting future direction would be examining possible nonbelieved memories in the Think/No Think paradigm and their related consequences in people's behavior.

Previous research has proposed that recollection has an adaptive value for humans in the sense that the flexibility of recollection can help humans survive in a complex and changing environment (Howe, 2011; Schacter, 2012). For example, after encountering exemplars of toxic insects such as *ant*, *mosquito*, *bee*, and so on, people will form false memories of prototypical toxic insects such as *spider* (e.g., Seamon, Luo, Schlegel, Greene, & Goldenberg, 2000). The false memory of a spider being a toxic insect can help people to avoid danger from a spider even though they have not actually encountered it before. However, the flexibility of recollection is much more limited than the flexibility of belief in occurrence. By incorporating information of past experiences, belief can be easily altered and this has an obvious functional advantage. As this dissertation shows, by quickly changing belief, human behavior can easily be adapted which might have survival relevance. Future research may want to investigate the adaptive value of belief versus recollection.

Another future step following this project would be to investigate the mechanisms underlying the belief effect. It is unknown why belief has such a profound impact on behavior while recollection does not. Although we have discussed possible mechanisms here, these propositions are more like speculations other than explanations. Recollection, or mental representation of a past experience, has always been considered as the core of remembering. The function and consequences of recollection found here raises more general questions such as the function of memory and why we need mental representations of our past (e.g., as an extreme example, one may merely have a set of beliefs validated by past experiences). Of course, those are far beyond the scope of this trivial dissertation. Further research is needed to investigate the mechanisms underlying the consequences of belief and recollection.

CONCLUSIONS

The goal of the present dissertation was to examine the cognitive-behavioral consequences of belief and recollection. The studies described here have – hopefully – convincingly shown that autobiographical belief of past experiences rather than recollection of past experiences is critical in behavioral performance. The idiom “seeing is believing” implies that one needs to experience something in order to believe the experience. Indeed, “seeing” (i.e., experiencing events) creates mental recollections of these experiences that are often accompanied by strong belief in the occurrence of those experiences. However, this dissertation suggests that merely “believing” in something will lead to similar consequences as actually seeing or experiencing an event. Once you believe an experience, it is not different from actually experiencing it in terms of the consequences on our behavior. Hence, the reverse version of the idiom – *believing is seeing* – can best summarize my conclusion: believing is as powerful as seeing in guiding our behavior.

References

Anderson, M. C. (2003). Rethinking interference theory: Executive control and the mechanisms of forgetting. *Journal of memory and language, 49*, 415–445.

Anderson, M. C., & Green, C. (2001). Suppressing unwanted memories by executive control. *Nature, 410*, 366–368.

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior, 22*, 261-295.

Berkowitz, S. R., Laney, C., Morris, E. K., Garry, M., & Loftus, E. F. (2008). Pluto behaving badly: False beliefs and their consequences. *The American Journal of Psychology, 121*, 643-660.

Bernstein, D. M., & Loftus, E. F. (2009). The consequences of false memories for food preferences and choices. *Perspectives on Psychological Science*, *4*, 135-139.

Bernstein, D. M., Laney, C., Morris, E. K., & Loftus, E. F. (2005). False beliefs about fattening foods can have healthy consequences. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 13724-13731.

Berntsen, D. (2010). The unbidden past: Involuntary auto-biographical memories as a basic mode of remembering. *Current Directions in Psychological Science*, *19*, 138–142.

Brogden, W. J. (1939). Sensory pre-conditioning. *Journal of Experimental Psychology, 25*, 323-332.

Brown, A. S., Caderao, K. C., Fields, L. M., & Marsh, E. J. (2015). Borrowing personal memories. *Applied Cognitive Psychology, 29*, 471-477.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Eichenbaum, H., & Cohen, N. J. (2004). From conditioning to conscious recollection: memory systems of the brain. New York: Oxford University Press.

Fazio, L. K., Barber, S. J., Rajaram, S., Ornstein, P. A., & Marsh, E. J. (2013). Creating illusions of knowledge: Learning errors that contradict prior knowledge. *Journal of Experimental Psychology: General*, *142*, 1-5.

Garcia, J., & Koelling, R. A. (1966). Relation of cue to consequence in avoidance learning. *Psychonomic science*, *4*, 123-124.

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.

Horowitz, M. (1986). Stress-response syndromes: A review of post-traumatic and adjustment disorders. *Hospital and Community Psychiatry*, *37*, 241-249.

Howe, M. L. (2011). The adaptive nature of memory and its illusions. *Current Directions in Psychological Science, 20*, 312-315.

Howe, M. L., Wimmer, M. C., Gagnon, N. & Plumpton, S. (2009). An associative activation theory of children's and adults' memory illusions. *Journal of Memory and Language*, *60*, 229–251.

James, W. (1890/1950). *The principles of psychology* (Vol. I). New York: Dover.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological bulletin*, *114*, 3-28.

Klein, S. B. (2011). *Learning: Principles and applications*. Sage Publications.

Lambert, A. J., Good, K. S., & Kirk, I. J. (2010). Testing the repression hypothesis: Effects of emotional valence on memory suppression in the think–no think task. *Consciousness and Cognition*, *19*, 281-293.

Lampinen, J. M., Meier, C. R., Arnal, J. D., & Leding, J. K. (2005). Compelling untruths: Content borrowing and vivid false memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 954-963.

Lampinen, J. M., Ryals, D. B., & Smith, K. (2008). Compelling untruths: The effect of retention interval on content borrowing and vivid false memories. *Memory*, *16*, 149-156.

Lewandowsky, S., Ecker, U. K., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, *13*, 106-131.

Luus, C. A., & Wells, G. L. (1994). The malleability of eyewitness confidence: Co-witness and perseverance effects. *Journal of Applied Psychology*, *79*, 714-723.

Mahr, J., & Csibra, G. (2018). Why do we remember? The communicative function of episodic memory. *Behavioral and Brain Sciences*, *41*: e1.

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

McCabe, D. P., Roediger III, H. L., McDaniel, M. A., & Balota, D. A. (2009). Aging reduces veridical remembering but increases false remembering: Neuropsychological test correlates of remember–know judgments. *Neuropsychologia*, *47*, 2164-2173.

Mitchell, K. J., & Johnson, M. K. (2000). Source monitoring: Attributing mental experiences. In *The Oxford handbook of memory*, pp. 179-195. Oxford: Oxford University Press.

Mitchell, K. J., & Johnson, M. K. (2009). Source monitoring 15 years later: what have we learned from fMRI about the neural mechanisms of source memory? *Psychological bulletin*, *135*, 638-677.

Morewedge, C. K., & Kahneman, D. (2010). Associative processes in intuitive judgment. *Trends in Cognitive Sciences, 14,* 435-440.

Otgaar, H., Moldoveanu, G., Wang, J., & Howe, M. L. (2017). Exploring the consequences of nonbelieved memories in the DRM paradigm. *Memory*, *25*, 922-933.

Otgaar, H., Muris, P., Howe, M. L., & Merckelbach, H. (2017). What drives false memories in psychopathology? A case for associative activation. *Clinical Psychological Science*, *5,* 1048-1069.

Roediger, H. L., Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review, 8*, 385-407.

Roediger, H. L., Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review*, *8*, 385-407.

Rubin, D. C. (2005). A basic-systems approach to autobiographical memory. *Current Directions in Psychological Science*, *14*, 79-83.

Rubin, D. C. (2006). The basic-systems model of episodic memory. *Perspectives on Psychological Science, 1,* 277–311.

Sauerland, M., & Sporer, S. L. (2009). Fast and confident: Postdicting eyewitness identification accuracy in a field study. *Journal of Experimental Psychology: Applied, 15*, 46–62.

Schacter, D. L. (2012). Adaptive constructive processes and the future of memory. *American Psychologist*, *67*, 603-613.

Schacter, D. L., & Addis, D. R. (2007). The cognitive neuroscience of constructive memory: remembering the past and imagining the future. *Philosophical Transactions of the Royal Society B: Biological Sciences, 362*, 773-786.

Schwarz, N., Sanna, L. J., Skurnik, I., & Yoon, C. (2007). Metacognitive experiences and the intricacies of setting people straight: Implications for debiasing and public information campaigns. *Advances in Experimental Social Psychology*, *39*, 127-161.

Scoboria, A., Boucher, C., & Mazzoni, G. (2015). Reasons for withdrawing belief in vivid autobiographical memories. *Memory*, *23*, 545-562.

Scoboria, A., Jackson, D., Talarico, J., Hanczakowski, M., Wysman, L., & Mazzoni, G. (2014). The role of belief in occurrence within autobiographical memory. *Journal of Experimental Psychology: General, 143*, 1242–1258.

Scoboria, A., Mazzoni , G. , Jarry, J. , & Bernstein, D. (2012). Personalized and not general suggestion produces false autobiographical memories and suggestion-consistent behavior. *Acta Psychologica*, *139*, 225–232.

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

Seamon, J. G., Luo, C. R., Schlegel, S. E., Greene, S. E., & Goldenberg, A. B. (2000). False Memory for Categorized Pictures and Words: The Category Associates Procedure for studying memory errors in children and adults. *Journal of Memory and Language, 42*, 120-146.

Tulving, E. (2002). Episodic memory: from mind to brain. *Annual Review of Psychology, 53*, 1-25.

Walther, E. (2002). Guilty by mere association: evaluative conditioning and the spreading attitude effect. *Journal of Personality and Social Psychology, 82*, 919-934.

Wimmer, G. E., & Shohamy, D. (2012). Preference by association: How memory mechanisms in the hippocampus bias decisions. *Science, 338*, 270-273.

Wright, D. B., Self, G., & Justice, C. (2000). Memory conformity: Exploring misinformation effects when presented by another person. *British Journal of Psychology, 91*, 189-202.

SUMMARY

Do you believe in your memories? Almost everyone would say yes. Do you act according to your memories? Most people would say yes (e.g., people having a memory of bad service in a restaurant will probably not visit that restaurant again). Do you have memories that you no longer believe? About one in five respondents would say yes. Do you act according to your nonbelieved memories? The answer here is largely unknown. The phenomenon of nonbelieved memories has raised a theoretical issue about the constructs of memory. Recent research has found that memory has two dimensions, namely *belief* in the occurrence of a past experience and mental *recollection* of the experience. What we know already is that memories accompanied with beliefs impact people’s behavior. What we do not know is which component of memory, belief or recollection, largely guides our behavior. The current dissertation aimed to answer this question.

This dissertation first reviewed consequences of memories in real life and studied the consequences of memories in the lab. **Chapter 2** reviewed cases in China and research on false memories showing that false memories of eyewitness might have serious legal consequences in that innocent people might be convicted. **Chapter 3** examined consequences of true and false memories in a lab-based perceptual task. It was found that memories of presentation of pictures and words made identification of related pictures more quickly.

Having established that memories indeed affect people’s behavior, Part II of this dissertation investigated whether it was belief or recollection of past experiences that impacts behavior. Belief refers to whether people believe an event occurred or not. Recollection refers to mentally re-perceiving an event accompanied with sensory details. One may behave merely because of a belief (e.g., believing one had been allergic to peanuts as a child) or because of a recollection (e.g., vividly remembering one’s allergic reactions to peanuts as a child), or both may be important in determining behavior. **Chapter 4** studied the priming effect of belief versus recollection on problem solving. Memories of words are found to facilitate problem solving performance on word puzzles, but it was unknown whether this facilitation was due to recollection or belief. The results showed that when beliefs for word memories were undermined, previously seeing the words no longer facilitated solving puzzles, suggesting the important role of belief in influencing behavior.

Memories are also found to impact decision making. **Chapter 5** tried to reduce people's beliefs in memories to see whether it would impact decision making to gain money. Participants first learned memory associations such as an arm picture always paired with a blue circle. Later they learned that the blue circle could win them money. Normally participants would choose the arm picture to win money as well since it was associated with the rewarded blue circle. However, it was found that when people were told that their memory associations were wrong (i.e., the arm picture did not pair with the blue circle), people no longer believed in their memory associations and they no longer chose pictures that were associated with rewarded circles to win money.

Chapter 6 examined which component of memory drove people's food preferences. Participants were falsely suggested that they had been sick after eating egg salad in childhood and they were guided to imagine this event. Some participants developed vivid recollections of being sick after eating egg salad, while some did not form recollections but believed it happened. The purpose was to see if only belief or recollection was sufficient to change people's preferences over egg salad. The results found that belief, but not recollection regarding the food aversive event, determined people's preferences for egg salad. Taken together, we found a *belief effect:* relative to recollection, belief is more critical in impacting people's behavior.

In Part III, the dissertation explored potential mechanisms of memory to explain the belief effect. **Chapter 7**'s results support the spreading activation theory that memory is a network consisting of nodes (e.g., *egg salad*, *sickness*) and associations among the nodes (e.g., *egg salad* led to *sickness*). Activation of one node such as *egg salad* will spread to nearby related nodes such as *sickness* and impact related behavior such as preference for egg salad. Combining all the results, I propose that belief acts as a reality monitoring mechanism that monitors the associations in the memory network, thus it monitors the direction of spreading activation in the end. As a result, change of belief greatly changes behavior.

In this dissertation, it was found that once people believed that they had been sick after eating egg salad, they behaved as if they truly experienced the food event. Or once people believed that a word was not presented or an association did not exist, they behaved as if they truly did not experience these events. I conclude that "believing" is like "seeing" in terms of impacting people's behavior.

中文简介

(Chinese Summary)

你相信你的记忆吗？几乎每个人都会回答相信。你会根据你的记忆行事吗？大多数人都认为会，例如不去光顾一家你记忆中服务很差的餐厅。 你有自己不再相信的记忆吗？ 大约五分之一的人会回答有。 你会依据你不再相信的记忆行事吗？ 答案却是未知的。 无信记忆（nonbelieved memory）现象引发了关于记忆的理论的新思考。什么是记忆？最近的研究发现，记忆存在两个维度，一是脑海中对过去的回想或回忆（recollection），二是对过去的信念（belief）。目前我们知道的是，伴随着信念的记忆会影响人们的行为。我们不确定的是，到底是记忆的哪个维度，信念还是回想会对行为产生影响。 本论文旨在回答这一问题。

论文的第一部分回顾了现实生活中以及实验室情境下记忆带来的影响。其中，**第二章**介绍了中国的一些刑事案件和梳理了错误记忆的最新研究成果，结果表明中国的部分冤假错案极有可能是由于目击证人的错误记忆造成的。**第三章**研究了正确记忆和错误记忆对一项认知任务的影响。第三章的研究成果表明，对看过的文字和图片的记忆有助于快速识别这些文字和图片。

在研究了记忆对人们行为的影响后，论文第二部分探究了究竟是对过去经历的信念还是对过去的回想能够决定一个人的行为。信念指的是人们是否相信一件事情发生过；回想则是指人们回忆事件时脑海中对各种细节的心理体验。人们可能对过去某一件事件只有信念，例如相信自己小时候曾吃花生过敏（但其实回忆不起来整个事件了）；或者对某一件事有着清晰、生动的回忆，例如清楚地记得当时吃花生后的各种过敏反应。这些信念或者回忆都可能导致他/她不再吃花生。**第四章**比较了信念和回忆对解决问题的作用。已有研究表明，对文字的记忆可以帮助人们猜字谜，但第四章的研究结果发现，当人们对看过的文字的信念下降时，他们猜出字谜的数量也随之下降。这表明信念对行为的重要影响。

记忆也会影响人们的决策。**第五章**实验中人们的记忆受到实验人员的挑战，然后看他们的赌钱决策是否会受影响。实验参与者首先形成一些记忆联结，例如一把椅子的照片总是出现一个蓝色的圆形图案旁边。然后他们通过实践学习到，蓝色的

圆形图案会赢钱。通常人们会认为这把椅子也会带来好运让他们赢钱，就像有些赌徒总喜欢戴某一条手链或穿某一个颜色的衣服。接下来，实验参与者的记忆受到挑战，他们被告知他们的记忆有误，总在蓝色圆形旁边出现的不是一把椅子而是一张床。结果发现，参与者的信念因受到挑战而改变，他们在打赌游戏中不再“迷信”和选择椅子来赢钱。

第六章检验了是信念还是回忆导致了人们的食物偏好。实验参与者被错误地引导和想象他们小时候吃鸡蛋沙拉后生病了。一些参与者产生了非常清晰的吃了鸡蛋沙拉后生病的回忆，另有一些参与者则没有引发任何回忆但相信这一童年事件发生过。实验的目的在于研究仅有信念或者仅有回忆是否足以改变人们对鸡蛋沙拉的喜好程度。结果发现，对过去的信念，而不是回忆，决定了人对食物的偏好。以上所有结果表明，我们发现了一种信念效应（belief effect）：相比回忆，信念在人的行为中起着更为关键的决定性作用。

论文的第三部分探讨了信念效应背后可能的机制。**第七章**的研究结果支持了激活扩散（spreading activation）理论：记忆犹如一张由不同的节点（如鸡蛋沙拉，生病症状等节点）织成的网，节点和节点之间有着联结（如，鸡蛋沙拉→生病）。一个节点的激活，将扩散到记忆网络中邻近的节点上，比如想到鸡蛋沙拉就会联想到生病，从而影响了对鸡蛋沙拉的食欲。结合本论文中的结果，笔者认为信念在记忆网络中起到了对记忆联结的监测功能，即控制和改变节点之间的联结，从而控制了激活的扩散方向。因此，信念的改变很大程度上改变了行为。

在这篇博士论文中，笔者发现，当人们相信他们曾吃鸡蛋沙拉生病后，他们的行为表现得好像生病事件真的发生过一样；当人们不再相信一些词的记忆或者某些记忆联结后，他们表现得就好像这些事真的没有发生过一样，对自己的回忆视而不见。笔者把这些现象总结为一句古语：信则有，不信则无。

VALORISATION ADDENDUM

Relevance

In the first part of this dissertation, I have examined the relationship between memory and behavior. Specifically, I have demonstrated that true and false memories can have both positive and negative behavioral consequences. This work is of relevance to the legal field, where eyewitness testimony relies to a large degree on human memory. As the first few chapters in this dissertation have demonstrated, interrogative and interview protocols can be made in accordance with how memory functions, e.g. by eliminating suggestive questions. Specific regulations and laws (especially in China) can be made to protect eyewitness memories from misinformation. For instance, there are no specific rules in regulating eyewitness identification processes in the Criminal Procedure Law of China. Data show that 97% of the identifications in China are not carried out by investigators who are blind to the specifics of the case, which changes the confidence that witnesses have in their identification (i.e., the post-identification feedback effect) and also increases the risk of misinformation being transmitted to the eyewitness. Specific rules should be made to prevent this, such as that the identification should be hosted by investigators who are blind to the case.

Furthermore, the current dissertation investigated the roles of two memory components-- autobiographical belief and recollection--in impacting behavior. I found that belief, not recollection, impacted various types of behavior such as problem solving, decision making and food preferences. Undoubtedly, this conclusion can shed light on situations where human memory is involved, including the legal field and the clinical field. Understanding and recognizing the role of autobiographical belief can help discriminate belief from recollection in eyewitness reports. The research results can also help develop therapeutic procedures aimed at altering beliefs in patients with trauma in order to ease their symptoms. For example, intrusive memory is a typical symptom experienced by patients with trauma. This dissertation points out possibilities that changing belief in trauma memory might affect beliefs in intrusive memories.

Target groups

This dissertation examined various behavioral and cognitive consequences of autobiographical belief and recollection, so the results of this dissertation are relevant to various groups of people. Chapter 2 reviewed Chinese cases of eyewitness misidentifications and how false memories may play a role in the Chinese legal system. Chinese policy makers and legal professionals including the police, lawyers, and judges may use this knowledge to understand the malleability of eyewitness memories, to design proper investigative and interview protocols that lower chances of memory distortions, and to make appropriate amendments (e.g., taking into account the reliability of memory when collecting eyewitness evidence) to certain articles in the Criminal Procedure Law in order to protect eyewitnesses' memories.

The results obtained in this dissertation have the potential to benefit patients and clinical psychologists as well. Chapters 4, 5 and 6 found that by changing one's belief (e.g., about a negative autobiographical event), people were no longer impacted by past events. For people who are suffering from trauma, this conclusion may be very relevant since beliefs in memories strongly impact cognitive-behavioral treatment outcomes. At present, there are certain therapeutic techniques such as imagery rescripting and Eye Movement Desensitization and Reprocessing (EMDR) that try to alter patients' memories to treat their symptoms. Clinical psychologists may aim to develop therapeutic methods that re-shape patients' beliefs about their past rather than recollections (which may remain intact) to more effectively treat patients with trauma- and stress-related disorders. Furthermore, this dissertation can be of interest to the general public. That is, they can learn that their memories are easily shaped by feedback from others such as friends and that this might affect the decisions they make.

Activities and Products

As summarized in Chapter 2, concrete activities can emerge based on our research findings. Workshops and seminars can be organized in China to educate legal professionals about how mechanisms of false memory can impact eyewitness reports and thus influence the investigative procedures and execution of justice. Moreover, since the Criminal Procedure Law of China does not specify the rules of regulating eyewitness identification processes such as lineup administration, additional rules (e.g., blind lineup) can be added based on my review findings. For the general legal system at the global scale, the differentiation between autobiographical belief and recollection should be made when evaluating eyewitness reports by asking eyewitnesses explicitly if they could recall specific recollections or they believe the occurrence of certain events.

New therapeutic methods to treat trauma may be produced based on the findings obtained from the current dissertation. In current therapies that treat trauma such as imagery re-scripting, memories are targeted to be changed by “altering” the unwanted content of negative past experiences. With the differentiation of autobiographical belief and recollection in memory, this dissertation has concluded that autobiographical belief is the key to change behavioral outcomes. Thus, therapeutic methods that target patients’ beliefs about their past could be developed from current research results.

Innovation

This is the first doctoral thesis comparing the cognitive and behavioral consequences of believing and remembering. Most of previous literature emphasized the importance of remembering the past, but my dissertation is the first to show the dominant role of believing the past when the past is used to guide people’s behavior. Legal professionals should be cautious when dealing with eyewitness reports. That is, more attention should be made on checking whether these reports are based on recollections or merely autobiographical beliefs. Beliefs are easy to manipulate while recollections are more resistant to false feedback. This dissertation also adds new value to the clinical field, opening the possibility to re-evaluate components of memory and how to target memory as a way of treating patients.

Schedule and Implementation

Within this PhD project, I have had close collaboration with researchers worldwide, especially with researchers in the United Kingdom and in China. I have presented work emerged from my PhD project at conferences in the Netherlands, Belgium, Hungary, and Canada. I have written an article targeted for the general public in the InMind magazine, to explain autobiographical belief and recollection and their relevance to real life. In the near future, we attempt to establish collaborations with legal professionals in China to organize seminars and workshops about the malleability of memory and its impact in the courtroom.

ACKNOWLEDGEMENT

Looking back at my PhD period, four years seem to have passed in a blink of an eye. Four years ago, I arrived in Maastricht in a cloudy day (as it always is) to start a new journey in my life, and I have never been dissatisfied by the kindness and openness offered by my supervisors, my colleagues and the people I met here in Maastricht. I want to thank a lot of people for supporting me to conduct all the work in this thesis.

First, my gratitude goes to my supervisors, Henry, Tom, Mark, and Harald. Henry, there are tons of words that I want to say to express my gratitude! The chance that we got to work with each other still looks incredible to me. Four and half years ago, I was an about-to-graduate master student seeking for opportunities to do my PhD. I accidently came across a paper written by you and then I thought, hmm this guy has done some interesting work and I want to do interesting work too! I checked your working place and it said Maastricht—never knew it before. But it did not matter. We were soon connected via email and with the help of Tom and Harald, we managed to finish all materials for a funding application within a week! Then we got funded. You were always so enthusiastic, kind, supportive and inspiring during my PhD! You always had time for me no matter how busy you were. You have invested tremendous time and effort in helping me to become a researcher. I still remember you saying in your email that the first draft of my first manuscript was okay although it was a disaster, but that gave me confidence. You patiently corrected my manuscript, which gave me guidance. I will never become a real researcher without your support! You also cared a lot about my life here since it was my first time to live in a country completely different from China. Thank you for your hospitality and kindness! I will always remember the Christmas dinner I had with your family.

Tom, it was really nice to have you as my supervisor. I first met you during our skype meeting four and half years ago. At that time, I was nervous and your humorous talking had made me relax a lot. Not being my daily supervisor, you have offered me a lot of help in designing my experiments and writing my manuscripts. You have taught me to be as critical as possible at my own experiment in the first place. “The fact that previous researchers have used this method does not mean anything. There are people jumping out of the bridge; but that does not mean you have to jump out of a bridge.”, this is your quote coming to my mind now! Thank you for all your kindness and support!

Mark, my research guide, thank you so much for developing fantastic theories about memory and leading me to interesting research questions. We did not have a lot of opportunities to meet in person since you live in a non-European country, but we had frequent email discussions that inspired me to formulate new research questions. As the only native English speaker in our team, you have also played a very critical role in finalizing our papers, for which I owe you a lot! Without your help, I can barely become a researcher who can write.

Harald, you always motivated me and provided insight into my research when we met. You asked divergent and creative questions that inspired new ideas. You have helped me with the funding for my Phd in the first place. Without your help, we may not get the funding. Harald, Mark, Tom, and Henry, thank you all again! I can never imagine a better supervision team than I have ever had.

I want to thank all my colleagues in the Forensic Psychology Section. I am rather an introverted person, but you have created an enjoyable working climate. I try to avoid giving specific names because every one of you has been so kind to a Chinese girl who never worked and lived abroad before. I also want to thank our secretaries and the instrumentation staff for their kind support. I would like to thank my paranymphs especially, Sanne and Yu, for their support during my defense. I also want to thank the first friend I have met in Maastricht, Qing Xu (徐青), for all the happy times that we have hung out together. Furthermore, I want to thank Prof. Chu Zhou (周楚老师) at Fudan University and Prof. Xiuyan Guo （郭秀艳老师） at East China Normal University for their support when conducting my research in China.

Finally, I want to thank my family most. Guilin, thank you so much for supporting me unconditionally in all these years! Words cannot express my gratefulness. Zhenzhen, my dear sister, thank you for helping with the cover and the invitation; they look great! Thank you for bringing joy to our parents when I was so far away from home! My dear parents, thank you for always standing by my side and supporting my choice. Although we are apart from each other for 10,000 kilometers, I can feel that you always care about me. Whenever I needed help, you lent me a firm hand. My dear daughter, Shanning, thank you for accepting and loving a Mommy who is a little bit workaholic. Mommy loves you so much!

我最需要感谢的是我的家人。桂林，感谢你多年来对我的无条件支持！任何语言都不足以表达我对你的感谢。我亲爱的妹妹，王真，谢谢你帮我制作了封面和邀请函！也感谢你在我不在爸妈身边的时候照顾他们，给他们带来慰藉。我亲爱的爸、妈，感谢你们总是支持我的选择，做我坚强的后盾。我们虽然相隔一万公里，但我能感受到你

们时刻都在关心着我。在我需要帮助时，你们毫不犹豫地竭尽所能帮助我。最后，致我最亲爱的女儿嫱宁，感谢你接纳有一点工作狂的妈妈，妈妈爱你！

PUBLICATIONS

Wang, J., Otgaar, H., Howe, M. L., & Zhou, C. (in press). A self-reference false memory effect in the DRM paradigm: Evidence from Eastern and Western samples. *Memory & Cognition*.

Wang, J., Otgaar, H., Howe, M. L., Lippe, F., & Smeets, T. (2018). The nature and consequences of false memories for visual stimuli. *Journal of Memory and Language,101*, 124-135.

Wang, J., Otgaar, H., Smeets, T., Howe, M. L., Merckelbach, H., & Zhou, C. (2018). Consequences of false memories in eyewitness testimony: A review and implications for Chinese legal practice. *Psychological Research on Urban Society, 1*, 12-25.

Wang, J., Otgaar, H., Howe, M. L., Smeets, T., Merckelbach, H., & Nahouli, Z. (2017). Undermining belief in false memories leads to less efficient problem-solving behavior. *Memory, 25*, 910-921.

Wang, J., Otgaar, H., Howe, M. L., Smeets, T., & Merckelbach, H. (2015). Nonbelieved memories: about what will occur when you do not believe in your memories. *In-Mind Magazine*, 11.

Otgaar, H., Moldoveanu, G., **Wang, J**., & Howe, M. L. (2017). Exploring the consequences of nonbelieved memories in the DRM paradigm. *Memory, 25*, 922-933.

Otgaar, H., Howe, M. L., Smeets, T., & **Wang, J**. (2016). Denial-Induced Forgetting: False Denials Undermine Memory, But External Denials Undermine Belief. *Journal of Applied Research in Memory and Cognition, 5*, 168-175.

Otgaar, H., Howe, M.L., Brackmann, N., & **Wang, J**. (2016). When children are the worst and best eyewitnesses: Factors behind the development of false memory. In R. Nash & J. Ost (Eds), *False and Distorted Memories* (pp. 23-38), Psychology Press.

Otgaar, H., Howe, M. L., Memon, A, & **Wang, J**. (2014). The development of differential mnemonic effects of false denials and forced confabulations. *Behavioral Science & The Law*, *32*, 718-731.

Otgaar, H., Howe, M.L., Clark, A., **Wang, J**., & Merckelbach, H. (2015). What if you went to the police and accused your uncle of abuse? Misunderstanding concerning the

benefits of memory distortion: A commentary on Fernandez (2015). *Consciousness and Cognition, 33*, 286-290.

Zhou, C., **Wang, J**., & Zhou, W. (2014). The self-reference effect in false memory: A robust facilitation of self. *Journal of Psychological Science, 5*, 1079-1083. (in Chinese)

Wang, J., Cao, H. & Zhang, Y. (2012). Cognitive clues can impact rat's pain affection. *Chinese Journal of Pain Medicine, 18*, 670-674. (in Chinese)

Manuscripts under Review

Wang, J., Otgaar, H., Smeets, T, Howe, M. L., & Zhou, C. (under review). Manipulating memory associations changes decision-making preferences in a preconditioning task.

Wang, J., Otgaar, H., Bisback, A., Howe, M. L., & Smeets, T. (under review). The consequences of implicit and explicit beliefs on food preferences.

Otgaar, H., **Wang, J**., Fränken, J., & Howe, M. L. (under revision). Believing does not equal remembering: The effects of social feedback and objective false evidence on belief in occurrence, belief in accuracy, and recollection.

ABOUT THE AUTHOR

Jianqin Wang (王俭勤) was born on 25 December, 1988 in Jinggangshan City, China. She received primary and high-school education in Jinggangshan City. In 2006, she was the champion (with the highest score) in the Entrance Examination for College in Jinggangshan City, and ranked around 200th among 350,000 participants in Jiangxi Province. She was admitted to Fudan University in Shanghai in 2006, studying Psychology. She received Renmin Scholarship and Jili Scholarship at Fudan University. During her undergraduate study, she wrote research proposals and gained two undergraduate-level grants for funding her research. She received a Challenged Scholar Award for one of her undergraduate research projects. Jianqin obtained her bachelor's degree with *cum laude* in 2010. After that, she had a gap year where she worked as an analyst for a commercial firm. Jianqin realized that research is her passion, so she got in a three-year master program of Psychology at Fudan University. She also managed to receive graduate student-level funds to support her research on memory. She obtained her master's degree with *cum laude* in 2014. From 2014 to 2018, she was a PhD candidate working on a project regarding the consequences of belief and recollection.