7-14-2020 4:00 PM

# The Conceptual Metaphor False Memory Effect

Jeffrey N. Reid, *The University of Western Ontario*

Supervisor: Albert Katz, *The University of Western Ontario*
Co-Supervisor: Stephen Lupker, *The University of Western Ontario*
A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Psychology

# Abstract

Conceptual Metaphor Theory posits that cross-domain mappings play a fundamental role in thought. However, to date there has been little research investigating the influence of conceptual metaphors in the subdomains of cognitive psychology, such as learning, concepts, and memory, leading critics to argue that conceptual metaphors are not psychologically real. The purpose of this dissertation was to explore whether conceptual metaphors influence episodic memory. In four experiments, a modified version of the Deese-Roediger-McDermott (DRM) paradigm was employed in which participants studied lists of expressions. Every expression within each list was based on a proposed conceptual metaphor. For example, the TIME IS MONEY list had expressions such as “how did you spend the summer break?”, “budget your hours,” and “is that worth your while?”. Following each list was a recognition test consisting of old (was on the list) and new (was not on the list) items. Critically, some of the new items were expressions that were based on the same conceptual metaphor as the study list (e.g., “that cost me a day”). Other new items were control expressions that talked about a similar topic but were not based on the same metaphor (e.g., “the weekend seems so far away”). In all four experiments, participants were more likely to falsely recognize new expressions that were metaphorically consistent with the study list than control expressions. These experiments demonstrate a clear influence of conceptual metaphors on memory, bolstering the claim that conceptual metaphors are psychologically real. Furthermore, it was found that participants showed the memory effect despite rarely reporting conscious awareness of the conceptual metaphors (Chapter 3). Participants also showed the effect when their attention was divided, which is known to diminish conscious and effortful processing (Chapters 4 and 5). Overall, these experiments provide converging evidence that conceptual metaphors are psychologically real and influence cognition automatically and unconsciously.

# Keywords

Conceptual Metaphor Theory, Deliberate Metaphor Theory, Metaphor, False Memory, Concept Representation, Divided Attention, Conscious vs. Unconscious Processes, Psycholinguistics.

# Summary for Lay Audience

Metaphors are ubiquitous in language, and much of the everyday language we use is actually metaphorical. For example, a phrase such as “I see your point,” when agreeing to an argument just made, is metaphorical because there is nothing physical to see. Rather, this expression is based on an underlying metaphor that UNDERSTANDING IS SEEING, as are other common expressions such as “look at the big picture” or “we have different views on this issue.” Beginning around the 1980’s, linguists began to consider metaphor not just as a special form of language, but as a fundamental component of thought, what they labeled “conceptual metaphors.” Because abstract concepts or ideas are not experienced directly, they are difficult to understand. Therefore, to understand these concepts we use conceptual metaphors that draw on concrete experiences, such as by comparing thoughts to our visual experiences.

Although linguists have made a compelling case that we use metaphors to think, the idea has not gained as much traction in psychology. Some psychologists argue that there is little experimental evidence that conceptual metaphors play a role in basic psychological phenomena such as problem-solving or memory. The purpose of this dissertation was to conduct psychological experiments to see if metaphor really does influence thought, and in particular, memory. In a series of experiments, I presented participants with lists of expressions that were all based on one underlying hypothesized “conceptual metaphor,” such as UNDERSTANDING IS SEEING. Following each list was a memory test in which several old (i.e., was on the list) and new items (i.e., was not on the list) were presented and participants had to identify the old items. Critically, some of the “new” items were based on the same conceptual metaphor as the old items. In each experiment, I found that participants falsely recognized these items; they thought these items had been presented before even though they were never on the study list. This finding demonstrates that metaphors influence how we remember information, in support of the argument that people use metaphors to think.

# Co-Authorship Statement

Chapter 2 of this dissertation is a manuscript accepted for publication that was co-authored by my supervisor, Dr. Albert Katz. I was the first (lead) author and wrote the majority of the "methods" and "results and discussion" sections. I also designed the study, recruited the research participants, and collected and analysed the data. Albert Katz wrote substantial portions of the introduction, edited the manuscript, and provided feedback on the study design and analysis.

# Acknowledgments

First and foremost, I give thanks to my Lord and Savior, Jesus Christ. Next, I would like to thank the best supervisor in the world, Albert Katz. Thank you for all your kindness and support. I'm very proud of the work we've done together, but I am even happier to call you my friend. Joining your lab was one of the best decisions I've ever made, and these past four years have been a blast! Thank you to Hamad Al-Azary. I couldn't have done this without you. You were someone I looked up to when I first started grad school and I really appreciate that you are always willing to lend a helping hand on anything from research ideas to grant applications and job applications. I really enjoyed all of our discussions on metaphor, and I look forward to all the projects we can work on in the future. Thank you to Huilan Yang. You are a great researcher and a great friend, and I hope we can do many more projects together. Thank you also to Sophie Liu for taking good care of me and supporting me emotionally during my PhD. I love all of you very, very much!

Thank you also to Steve Lupker for joining as a co-supervisor on this thesis after Albert retired. It was very nice to get to know you and your lab members over the last several years.

I also want to thank all of the examiners on my thesis, both at the departmental level and at the senate level, which includes Herbert Colston, Deanna Friesen, Marc Joanisse, Ken McRae, and Paul Minda. All of you brought valuable insights to my research, and I greatly enjoyed the discussions.

I also would like to thank all of my friends who helped me during my PhD. Thank you to Alex Taikh, Krysten Zator, Karen Hussey, James Boylan, Giacomo Spinelli, Marie Ostia, Tina Yu, Sydney Rossiter, Miranda Qi, Mark McPhedran, Lingling Li, Zian Chi, and Angela Zhou. I hope that our paths will continue to cross in the future!

Lastly, thank you to my family back home: Mom, Dad, Kristyn, and Katie. Thank you for all the long-distance support.

As is appropriate for this circumstance, I will end with a metaphorical expression that instantiates the BODY IS A CONTAINER FOR EMOTIONS conceptual metaphor: *my heart is overflowing with gratitude*. Thank you all very, very much!

# Table of Contents

# List of Tables

# List of Appendices

# Chapter 1

# 1 Introduction

Conceptual Metaphor Theory (CMT) is an important theory in a wide range of disciplines, with Lakoff and Johnson's seminal work, *Metaphors we Live By* (1980), being cited in over 63,000 publications (Google Scholar, retrieved May 18, 2020). Surprisingly, although CMT posits cognitive representations and processes, it has had little impact in cognitive psychology. The main aim of the following studies is to provide direct tests of some of the fundamental assumptions of CMT using the Deese/Roediger-McDermott (DRM) paradigm (Roediger & McDermott, 1995). The first goal is to test whether conceptual metaphors are psychologically real in the sense they produce a false memory effect, as would be expected when studied using an episodic memory test. A secondary goal is to explore whether conceptual metaphors, if they do show memory effects, operate automatically or require conscious and deliberate attention.

## 1.1 Conceptual Metaphor Theory

The theory posits that the human conceptual system is metaphorical in nature, such that knowledge is mapped from one conceptual domain onto a second conceptual domain, via "conceptual metaphors." These metaphors can be inferred from examination of linguistic, often metaphorical, expressions. Indeed, the bulk of support for the existence of conceptual metaphors comes from text exegesis. For instance, consider the ostensibly unrelated expressions "*that cost me a day,*" "*budget your hours,*" and "*how did you spend the summer?*" Lakoff and Johnson (1980) argued that these expressions are in fact not unrelated but bear an underlying similarity, namely that each of these expressions talk about a target domain (e.g., time: days, hours, seasons) in terms of a very specific source domain (e.g., money: cost, budget, spend). They argued further that this particular conceptual metaphor, TIME IS MONEY, partially structures one of the ways in which we understand the abstract concept of "time," and the inferences we derive about time. Inferences would include thinking about "time" as being a valuable and limited resource that needs to be managed wisely or else it will be wasted. According to Lakoff and Johnson, these inferences would not be possible without the conceptual metaphor.

Lakoff and Johnson (1980) propose that metaphorical expressions are understood by accessing the broader conceptual metaphor mapping upon which the expression is supposedly based. Importantly, the theoretical claim is that, on encountering an expression such as *"how did you spend the summer?*" the underlying conceptual metaphor mapping (TIME IS MONEY) is activated automatically and arouses a set of correspondences between the two separate mapped domains (Lakoff, 1993, 2008).

Consider an expression such as "*our relationship is at a dead-end.*" In this instance the theorized set of correspondences that is assumed to be activated relates our stored knowledge about journeys to the concept of love. When one considers a real or literal journey, one considers the travelers, the route, the destination, the mode of transport, and so on. When JOURNEY is mapped metaphorically onto the domain of LOVE, the "travelers" partaking in a literal journey correspond to the lovers, the "vehicle" corresponds to the relationship, and the "landmarks" found in a physical travel correspond to life-events found in relationships, such as a first date, a first kiss, a marriage, and so on (Katz & Taylor, 2008). Thus, for an expression such as "*we're at a crossroads,*" the correspondence would be that in a literal journey one must make a decision regarding which path one must take, and if there is only one vehicle, some travelers may have to leave the vehicle if they want to go in a different direction. When this information is mapped onto the LOVE domain, the meaning is that the lovers need to make a decision regarding whether they will continue together, and if one of the lovers wants to go in a different direction, they will have to leave the relationship. According to Lakoff (2008), correspondences such as these are automatically accessed upon encountering a triggering expression.

As Murphy (1996) has pointed out, because CMT is a theory of conceptual knowledge, there should be observable effects across many domains of cognition, such as memory, problem solving, learning, and categorization. Moreover, if CMT is a psychologically real theory then several testable hypotheses should follow from its assumptions, based on what is known about semantic memory and episodic memory-based research. Here I test implications with observations from an episodic memory task. These hypotheses include, first, if a metaphorical expression activates a conceptual metaphor, other expressions that are also derived from the same conceptual metaphor should also become partially activated. This follows from the spreading activation account of semantic memory, which proposes that when one concept (e.g., “mug”) is activated, other related concepts become partially activated as well (e.g., “cup,” “coffee,” “drink,” etc.). If a conceptual metaphor is psychologically real, then there should be analogous processing when a triggering metaphoric expression is encountered: other metaphoric instantiations of the conceptual metaphor also should be partially brought to mind. Second, but also crucially, expressions that come from a different conceptual metaphor, even if they are on a similar topic, should *not* be activated, or not highly activated. That is, an expression such as “*our relationship is at a dead-end*” should activate the LOVE IS A JOURNEY conceptual metaphor and related expressions (e.g., “*we’ve come a long way together*”) but should not activate, at least not to the same degree, the LOVE IS MAGIC conceptual metaphor and related expressions (e.g., “*she cast a spell on him,*” “*she’s very charming*”). Although in this example, the expressions all share the common element of LOVE, according to CMT the expressions are derived from distinct mappings. Therefore, the spread of activation from LOVE IS A JOURNEY expressions to LOVE IS MAGIC expressions should not be as strong as it would be to other LOVE IS A JOURNEY expressions. If the spread of activation was equal among all of these expressions, CMT would have little explanatory power as a psychological theory (Katz & Reid, 2020).

## 1.2 Previous experimental tests of CMT.

Research that has directly tested CMT’s assumptions has provided mixed support. For example, one might expect priming effects, a phenomenon well studied in the semantic memory literature.  Allbritton, McKoon, and Gerrig (1995) found support for CMT in a recognition priming experiment. Participants were presented with short paragraphs in which one of the early sentences instantiated a conceptual metaphor (e.g., CRIME IS A VIRUS). The next several sentences either continued talking about the TARGET domain (“congruent condition”) or shifted to talking literally about the SOURCE domain (“incongruent condition”). For instance, for the CRIME IS A VIRUS paragraph, the congruent condition continued to talk about a crime outbreak and extended the metaphor by mentioning that crime was starting to “infect” safe neighborhoods. In the incongruent condition, the paragraph instead mentioned that the police force had literally been infected with viral pneumonia. The final sentence, “*Public officials desperately looked for a cure,”* which was the same in both conditions, therefore either referred metaphorically to a cure for crime, further instantiating the CRIME IS A VIRUS mapping, or referred literally to a cure for pneumonia, not instantiating the mapping given the context. On a subsequent recognition test, participants were tested on the early sentence that first instantiated the conceptual metaphor and on the final sentence, in that order. Recognition times for the final sentence were faster in the congruent condition, suggesting that recognition of the early sentence primed recognition of the final sentence to a greater degree when both sentences instantiated the CRIME IS A VIRUS conceptual metaphor. Similar effects were found with single word recognition (e.g., crime-CURE). Although these results are consistent with CMT, some have argued that the findings do not necessitate that conceptual metaphors were activated automatically as proposed by CMT but could be explained if participants engaged in a deliberate memory strategy (Keysar, Shen, Glucksberg, & Horton, 2000; McGlone, 2007).

Studies that focus on online sentence reading also provide mixed evidence for CMT. These studies follow the same logic as Allbritton et al. (1995), but with online reading measures rather than recognition latencies. Participants first read short paragraphs in which a conceptual metaphor is instantiated by several conventional expressions. The final sentence, on which reading time is measured, is either based on the same or a different conceptual metaphor. Thibodeau and Durgin (2008) and Gong and Ahrens (2007) found priming effects supportive of CMT, but Glucksberg, Brown, and McGlone (1993) and Keysar et al. (2000) did not. Thibodeau and Durgin (2008) criticized the stimuli employed by Keysar et al. (2000), arguing that their conventional expressions were not really conventional and that the conditions were not properly matched in terms of how well the lead-up sentences conceptually aligned with the target sentence. However, Thibodeau and Durgin (2008) point out that their own effects could be attributed to lexical priming and not necessarily the consistency of the conceptual metaphor mapping, a position also argued by McGlone (2011). Although this was a potential confound, lexical priming typically has limited effects on reading natural discourse, especially in terms of total reading time for sentences (Hyönä, 1993; Traxler, Foss, Seely, Kaup, & Morris, 2000) which was the measure used by Thibodeau and Durgin. Therefore, lexical priming is unlikely to have compromised their conclusions.

As one further complication, Gong and Ahrens (2007) found that the presentation format of the sentences can also influence the results. Thibodeau and Durgin (2008) used a paragraph presentation format whereas Glucksberg et al. (1993) and Keysar et al. (2000) used line-by-line presentation. Gong and Ahrens argue that line-by-line presentation leads participants to expect new information, which hinders them from using the conceptual metaphor as a schema to process the paragraph. In a series of experiments, they found priming effects supportive of CMT when the stimuli were presented in paragraph form, but null effects when presented line-by-line. However, this is an *ad hoc* explanation and it seems unlikely that something as inconsequential as the presentation format could disrupt conceptual metaphor processing.

Although these studies suggest that conceptual metaphor priming is possible, it appears to only occur under certain conditions. Furthermore, Gong and Ahrens (2007) argue that participants' expectations influenced the results. If this is true, it suggests that priming depends on the participant consciously attending to the metaphorical language and noticing how it is repeated through the discourse. If conceptual metaphor activation is automatic, expecting new information should not matter as automatic processes are not easily disrupted. Therefore, the fact that the form of presentation did have an influence on reading times suggests that participants may have deliberately attended to the conceptual metaphor mappings.

Recall that the first aim of my thesis studies is to test whether false memory effects consistent with conceptual metaphors can be produced using the DRM paradigm. To anticipate slightly, the answer is "yes". A second aim of the proposed studies is to examine whether false memory effects one could attribute to CMT are due to automatic activation or deliberate conscious processing.

## 1.3 Deliberate Metaphor Theory (DMT)

Steen (2009) has proposed that expressions only activate a cross-domain mapping when they contain a “deliberate” metaphor. Steen (2008, 2009, 2011, 2015, 2017) defines a deliberate metaphor as one that is purposely used as a metaphor in that the speaker or writer intends the hearer or reader to see one domain in terms of another domain. Deliberate metaphors are identified by examining the language itself. For instance, metaphor is used deliberately if the statement or passage includes a direct comparison as in the classic Shakespearian metaphor: “*Shall I compare thee to a summer’s day.*” Steen (2008, 2009) argues that this can be considered a deliberate metaphor because the author explicitly mentions that a comparison is taking place (“shall I compare”) and the reader cannot help but see the individual being described in terms of an alien domain (i.e., a summer’s day). Other indicators of deliberate metaphor use are pragmatic markers (lexical items that signal a comparison, such as “one might say”; Gibbs, 2015; Steen, 2008), extended comparisons in which several elements of one domain are mapped onto the other domain (Steen, 2008, 2015), novel extensions of conventional metaphors (Steen, 2009), similes, in which the comparison is also explicitly stated (Steen, 2009), and rejections of conventional metaphors (e.g., a therapist asking a client to think of the harmful effects of framing cancer as a war; Steen, 2011). Researchers of deliberate metaphor do not directly ask the speaker or writer whether they intended to use a metaphor or not. Instead, Reijnierse, Burgers, Krennmayr, and Steen (2018) have a procedure for identifying deliberate metaphor use in text and it is assumed that if certain features such as those mentioned above are present, then the metaphor is being used deliberately.

According to Steen (2009), only deliberate metaphors afford conscious metaphorical thought, and conceptual metaphor mappings only influence cognition (at least in the ways proposed by CMT) when the participant is consciously aware that a metaphor is being used. As he puts it: “Contrary to what CMT assumes, the power of metaphor may not lie in its widespread unconscious use but in its much more restricted and targeted deliberate – sometimes conscious – use” (2009, p. 194). In other words, for conscious metaphorical thought to occur, the metaphor must be deliberate (and even deliberate metaphors do not always evoke conscious metaphorical thought), and the conceptual mapping is only engaged when the reader (or hearer) consciously realizes that it is a metaphor. If the reader does not consciously realize a metaphor is being used, they may process the metaphor using lexical disambiguation. For instance, for a conventional metaphor such as "how did you *spend* the weekend?", Steen argues that the hearer would simply access the alternate meaning of *spend* related to "devoting time" in his or her lexicon rather than activating a cross-domain mapping between TIME and MONEY. In this case, the metaphorical expression should not activate other expressions based on the same conceptual metaphor because the conceptual mapping itself is not activated.

Therefore, Steen is critical of CMT's claim (see Lakoff, 1993) that conceptual metaphors are engaged automatically and unconsciously. Although Steen has the most elaborated critique of automatic conceptual metaphor activation, others have also argued that psychological findings seemingly supportive of CMT may actually be due to participants consciously attending to the metaphors rather than the mappings being activated automatically and unconsciously (Glucksberg et al., 1993; Glucksberg & McGlone, 1999; Keysar, Shen, Glucksberg, & Horton, 2000; McGlone, 2007).

Although deliberateness may be an interesting variable, especially its role in communication, the more relevant claim of that position for the current thesis is that one needs to be consciously aware that metaphors are being used for conceptual metaphors to influence his of her cognition. Rather than manipulate deliberateness, attention will be manipulated here to directly assess whether cognitive effects of conceptual metaphors depend on the participant's conscious awareness of the metaphors being used. Indeed, if conscious awareness is not required, then the question of whether only deliberate metaphors evoke conscious awareness of metaphoricity may no longer be relevant, at least regarding how conceptual metaphors affect cognition.

## 1.4 The current research

The research reviewed above demonstrates that results supportive of CMT can be found using cognitive measures, such as reaction time, at least under certain conditions, though the question of automaticity or deliberateness of these effects are to be determined. As noted above, CMT is a broad theory of semantic knowledge, and as such, conceptual metaphors, if they exist, should inform all domains of cognition. The research reported in this thesis focused on the effects of conceptual metaphors in a specific episodic memory task, namely a variant of the Deese-Roediger/McDermott (DRM; Roediger & McDermott, 1995) task. In this task, one finds false memory for non-presented items related to the items on a study list. As extended to the test of CMT, I argue that analogous effects should obtain if the study list consists of expressions that, presumably, draw on a common conceptual metaphor. The theory then should lead to the prediction that false memories should be observed for items not presented at study but that are also drawn from the same conceptual metaphor.

### 1.4.1 Use of episodic memory to study CMT

Episodic memory involves memory for discrete events connected to a time and place in a person's past (Tulving, 1972). Although the participant may not be able to articulate exactly where or when the event occurred, they remember the occurrence of the event and its temporal-spatial relation with other events (e.g., I remember going for coffee after seeing that movie). In contrast, semantic memory involves knowledge about the world, such as facts and meanings of words, but one does not typically remember when and where this knowledge was learned. A task such as the DRM, in which the participant must indicate whether a word occurred in the experiment, is episodic because it asks whether an event (i.e., the word) occurred at a specific place and time (i.e., during the experiment). In contrast, a question such as "what is the capital city of Ontario?" would involve semantic memory; that is, retrieving knowledge not connected to a discrete event in the person's past. Although the person would have had to learn this fact at a specific time and place, this temporal-spatial information no longer needs to be connected to this fact for the question to be answered. Although there is debate over how distinct the two memory systems are (Renoult, Irish, Moscovitch, & Rugg, 2019; Tulving, 2002), it is clear that stored knowledge of the world influences how one organizes the encoding and retrieval of experienced episodic information, such as read verbal material (Bousfield, 1953; Tulving & Thomson, 1973, Weidemann et al., 2019).

To my knowledge, only two studies have explicitly employed episodic memory tasks to examine CMT (Katz & Law, 2010; Reid & Katz, 2018a). Katz and Law (2010) employed the release from proactive interference (PI) paradigm, a task in which participants are given consecutive short lists of items to remember and the items on all the lists are related on some dimension (e.g., members of the same taxonomic category). Typically, a free recall test follows each list. Many studies have found a decline in recall over lists (see Wickens, 1970, for a review) with one of the main explanations being that the retrieval cue (e.g., the category label) becomes overloaded, and thus, ineffective (Gardiner, Craik, and Birtwistle, 1972; Watkins & Watkins, 1975). This effect is called the buildup of proactive interference as previously learned items from early lists interfere with recall of subsequent lists. A "release" from PI occurs when the items change on a dimension, for instance, if a list contains items from a different taxonomic category from the previous lists. When this change occurs, recall typically improves and sometimes fully recovers from the earlier decline due to build-up of PI. Katz and Law (2010) hypothesized that a similar build-up and release of PI would occur with metaphor expressions based on conceptual metaphors, that is, as consecutive lists containing expressions from the same conceptual metaphor are presented, recall would decline, but if a later list contained expressions from a different conceptual metaphor, recall would recover. The theory behind this was that the conceptual metaphor would act as a retrieval cue that overloads as more exemplars are presented. When the conceptual metaphor changes, it acts as a new, effective retrieval cue. Katz and Law's hypothesis was confirmed. They observed an initial decline in recall when the lists all contained expressions from the conceptual metaphor LIFE IS A JOURNEY, followed by an increase in recall when the final list changed to LOVE IS MAGIC expressions. Although this supported CMT, the results are limited because the topics of the sentences also changed, that is, the build-up and release from PI could be due to the sentences all being about LIFE (or even JOURNEYS) initially and changing to sentences about LOVE (or MAGIC). In other words, the effects could be due to topical similarity, and not due to metaphorical mappings. As will be seen, I controlled for this in the DRM paradigm by including control lures on the recognition tests that shared either the same target or source domain as the expressions in the presented study list but did not use the same metaphorical mapping.

The only other published study testing CMT using an episodic memory task of which I am aware is the DRM study by Reid and Katz (2018a). This study has been published and, because it is the foundation for all subsequent studies in this thesis, is included here as Chapter 2.

### 1.4.2 Employing the DRM procedure

The DRM procedure is a well-known episodic memory task influenced by semantic information, and therefore, affords an explicit test of CMT. In the standard DRM paradigm, participants are given multiple lists of words, with all items in a list associated to one non-presented concept (Roediger & McDermott, 1995). For instance, the list might include a number of words that are associated to "sleep", such as *bed, awake, rest, doze,* and *pillow*. However, the word "sleep" is not presented during the study phase. After processing the study list, participants are asked to recall the list or to recognize the presented items from a set that also contains "lures." The lures consist of new items not associated with the list at all and the critical lure, the item around which the study list was constructed but was not presented in the study phase. Typically, participants falsely remember the non-presented critical lure, sometimes even as often as presented words.

One of the popular theories to explain this effect is based on the postulation that when one concept is activated in semantic memory, activation spreads to other semantically related concepts, known as "spreading activation" theory (e.g., Roediger, Balota, & Watson, 2001). With the DRM paradigm, each of the presented concepts partially activates the critical non-presented concept, and if this concept receives enough activation across all the presented words then it could induce a false memory in a later episodic memory test. One of the explanations for false memory that goes hand in hand with spreading activation is "processing fluency," which refers to the ease with which stimuli are processed (Jacoby & Dallas, 1981; Jacoby & Whitehouse, 1989; Kelley & Jacoby, 1998). There are two types of fluency: perceptual and conceptual. Perceptual fluency occurs when a stimulus is more easily processed perceptually, such as if the stimulus is flashed on the screen for a short duration in lowercase letters before being shown in uppercase levels for the recognition judgment (e.g., mug-MUG; Jacoby & Whitehouse, 1989; Rajaram, 1993).

Conceptual fluency involves a semantic associate (e.g., drink-MUG) being flashed on the screen, leading to semantic, but not perceptual priming of the to-be-recognized stimulus (Rajaram & Geraci, 2000). Conceptual fluency is thus a by-product that occurs when associated concepts are primed through spreading activation, leading to more fluent processing in a subsequent recognition test. Fluency in general affects recognition. Jacoby and colleagues (Jacoby & Dallas, 1981; Jacoby & Whitehouse, 1989; Kelley & Jacoby, 1998) argue that fluency is a heuristic used to judge the familiarity of a stimulus. This not only influences correct recognition, but also false recognition because the participant mistakes the ease with which they process the stimulus as familiarity with it (Jacoby & Whitehouse, 1989). Both conceptual (Rajaram & Geraci, 2000) and perceptual (Jacoby & Whitehouse, 1989; Rajaram, 1993) fluency increase false recognition.

Conceptual fluency has been proposed as a possible mechanism that causes or at least contributes to the DRM false memory effect (Doss, Bluestone, & Gallo, 2016; Gallo & Roediger, 2003; Whittlesea, 2002). Whittlesea found that after reading a DRM-type list of associates, the critical lure was processed faster in a subsequent lexical decision task. This suggests that the study list enhanced semantic processing of the lure. Therefore, DRM study lists may parallel semantic masked priming manipulations at recognition (Rajaram & Geraci, 2000) as both lead to conceptually fluent processing of the lure.[1] This fluency may then be (mistakenly) attributed to familiarity with the item (Jacoby, Kelley, & Dywan, 1989).[2]

According to CMT, when one encounters a metaphorical expression, the underlying conceptual metaphor is automatically activated, and this activates the entire set of

---

[1] Note that increased processing fluency does not necessarily lead to faster recognition response latencies (Rajaram, 1993; Rajaram & Geraci, 2000). Unlike a lexical decision task that more directly measures processing speed, recognition is a judgment of whether the stimulus was previously experienced. Therefore, in a recognition task, after the stimulus has been processed the participant must still make a memory decision. As such, multiple factors contribute to a recognition decision and recognition latency is not a pure measure of processing speed.

[2] Although the focus here is on false recognition, see Leboe and Whittlesea (2002) for how an attribution process could underlie false recall as well.

correspondences between the two domains. If the entire set of correspondences is indeed activated, presumably other expressions from the same conceptual metaphor should be processed more fluently. Therefore, like other conceptual fluency manipulations, conceptual metaphors should induce false recognition.

The other major theory of false memory is fuzzy-trace theory, which proposes that participants extract the meaning, or "gist" of the stimuli they are presented, and make memory errors because non-presented items are consistent with the gist of the presented items (Reyna & Brainerd, 1995; Brainerd & Reyna, 2002). One can in principle extend this theory to how conceptual metaphors may induce false memories. When a participant reads a list of expressions based on one underlying metaphor mapping, this mapping could be considered the "gist" of the list. The conceptual metaphor itself, as well as other non-presented expressions based on the same mapping would be consistent with the gist, and thus, should be likely to be falsely remembered. Unlike spreading activation which posits automatic activation, fuzzy-trace theory does not take a stance on whether gist extraction is automatic or more intentional in nature (Brainerd, Forrest, Karibian, & Reyna, 2006). Therefore, if false recognition occurs with conceptual metaphor expressions, it does not necessarily prove that conceptual metaphors are activated automatically. Nonetheless, according to both spreading activation and fuzzy-trace theory, it is possible for conceptual metaphors to induce false memories.

The following chapters will detail four experiments in which I adapted the DRM paradigm for testing CMT by constructing lists of expressions based on conceptual metaphors. The hypothesis was that activation from reading the study list expressions would spread to other non-presented expressions that are based on the same conceptual metaphor mapping. As a result, these *non-presented* expressions would be more likely falsely recognized than control expressions that do not engage the same conceptual metaphor. Also, because each presented expression should activate the conceptual metaphor, the conceptual metaphor label itself (e.g., "*time is money*") may also be falsely recognized.

Chapter 2 is a published article in which we found evidence of a “conceptual metaphor false memory effect” using a simple old/new recognition test. Chapter 3 replicates and extends the study by assessing participants’ subjective experiences of false recognition and their strategies for remembering the lists. Chapters 4 and 5 directly assess the automaticity of conceptual metaphor activation by replicating the false memory effect under divided attention conditions at both study and test. Finally, Chapter 6 discusses the findings from all four experiments and implications for memory theories, CMT, and DMT.

# Chapter 2

## 2 Something false about conceptual metaphors

This is an **Accepted Manuscript** of an article published by Taylor & Francis in ***Metaphor and Symbol*** on **January 3, 2018**, available online at the Taylor & Francis Ltd web site: www.tandfonline.com

Link to article: https://www.tandfonline.com/doi/full/10.1080/10926488.2018.1407994

See appendix A for copyright permission letter.

Article Title: Something false about conceptual metaphors

Authors: J. Nick Reid and Albert N. Katz

Journal Information: *Metaphor and Symbol*, *33*(1), 36-47.

# Abstract

Although Lakoff and Johnson's Conceptual Metaphor Theory has been influential across many disciplines, little research has tested the psychological reality of conceptual metaphors (CMs) using established experimental memory paradigms. Here we employ an episodic memory task based on the DRM false memory paradigm to explore this possibility. We find that after reading lists of sentences based on underlying conceptual metaphors that participants are more likely to falsely remember the non-presented conceptual metaphors themselves as well as new sentences consistent with the CM mapping than control items that do not share this mapping. This finding provides experimental support for conceptual metaphors and highlights the utility of using episodic tasks to explore the assumptions of this theory.

## 2.1 Introduction

Metaphor has been considered traditionally as a matter of language and not as thought. Thus, on encountering a metaphor such as “my life is a dead-end street”, traditional comprehension models argue that an interpretation depends on mapping semantic properties of the concept “life” onto properties of the concept “dead-end street” (eg., Bowdle & Gentner, 2005) or in categorizing “dead-end streets” into an *ad hoc* category, such as “places that lead nowhere” (e.g., Glucksberg & Keysar, 1990). However, starting with Lakoff and Johnson’s seminal work, *Metaphors we Live By* (1980) the argument has been made that metaphor is fundamentally a matter of thought, and not merely language. This approach has become known as Conceptual Metaphor Theory (CMT). From the CMT perspective a sentence such as “my life is a dead-end street” is a metaphoric expression that is motivated and understood through an underlying metaphoric conceptual system, in which a concept (usually abstract), such as “love”, is mapped onto a target concept such as “journey”, based on experiential and embodied interactions with one’s environment. The LOVE IS A JOURNEY conceptual structure thus motivates and structures our understanding of love through our understanding of journeys, such as they have a beginning, an end, a route taken, possible impediments and detours, dead-ends and the like. From this perspective seemingly unrelated metaphoric expressions can be understood as sharing a more basic conceptual underpinning.

Consider, for example, the abstract concept of “time”. According to the CMT, our conceptualization of “time” is structured around the notion that time is like “money” a resource that is limited and can be depleted, implicating the conceptual mapping TIME IS MONEY. For example, we talk about time as if it can be spent (“I’ve been spending a lot of quality minutes with her lately”), wasted (“this project is a waste of time”), invested (“studying psychology is a good investment of your time”), hoarded (e.g., “I have more than enough months with which to complete that project”), and so on.

One cannot underestimate the intra-disciplinary influence of conceptual metaphor theory. It has become a dominant approach in a large number of disciples (see for instance the complete issue of the journal *Cognitive Semiotics* which not only demonstrates the range of applications, but the modifications, limitations and the promissory notes held by the theory, Fusaroli & Morgagni, 2013). Indeed the seminal work, “Metaphors we live by” has been cited almost 50, 000 times according to Google Scholar, and a later work by Lakoff (*The contemporary theory of metaphor*, 1993) over 6,000 times. One reason for the popularity of Lakoff and Johnson’s (1980) theory is not only its applicability to the many disciplines based on analyzing discourse but because it is an early and important contributor to broader theories of grounded cognition- a class of theories that argue cognition is fundamentally tied to perception and action, and that concepts are not merely amodal symbolic representations (e.g., Barsalou, 1999, 2008; Gibbs, 2006). One of the major issues with such theories is in offering an account of how abstract concepts could be represented in an embodied format. Conceptual metaphor theory offers a compelling solution for this, that is, that abstract concepts are understood metaphorically via more concrete, experiential domains.

Despite the strong adoption of CMT in many disciplines, and although Lakoff and Johnson’s (1980) theory makes many claims about concept representation, semantic memory, and language comprehension, the discipline in which it has failed to gain significant traction is in the one that most directly and experimentally studies concept representation, memory and language, namely cognitive psychology. There are several reasons for this lack of support (see Gibbs, 2009), though most frequently mentioned are those that revolve around claims that the theory is too underspecified for experimental testing and that supportive findings can be more parsimoniously explained by more accepted mechanisms, such as those based on associative mechanisms (McGlone, 2011). The paucity of convincing experimental support stand in contrast to the bulk of a supportive literature which is based largely on text exegesis, a methodology not favoured in experimental cognitive psychology. Additionally, the CMT is based on a hypothetical conceptual system that does not fit in easily with current models of online language comprehension or semantic memory found in the experimental cognitive science literature. Finally, the supportive evidence is often adduced from studies indicating the embodiment of concepts, such that, for instance, showing that physical closeness is related to ratings of similarity as suggested by a conceptual metaphor SIMILARTY IS CLOSENESS (e.g., Casasanto, 2008) or data indicating brain activity in areas associated with motor or sensory regions when one is processing linguistic information that referenced motor or spatial information (Desai, Binder, Conant, Mano, & Seidenberg, 2011). Although the findings from such studies are exciting and suggest that conceptual metaphors are experiential and grounded in embodied mechanisms, they nonetheless are inferential and subject to the criticism that one can show embodied cognition without necessarily invoking the presence of conceptual metaphors (see Barsalou & Wiemer-Hastings, 2005).

The aim of the paper presented here is to adopt a paradigm found in the experimental literature on memory and, by adapting a logic similar to that accepted in the memory literature, test to see if conceptual metaphors influence memory in the ways predicted by the extant literature. Specifically, in the studies reported here we employ methodology well-established in the memory literature in which it has been assumed that underlying semantic structures have been activated during the encoding of verbal materials and the results of this activation are present during a later memory test of the originally presented materials. We also aim to test whether an alternative explanation, based on associative similarity, could explain the effects observed.

It should be noted that there is a very limited literature that has employed memory tasks to test CMT. Katz and Taylor (2008) examined the psychological reality of conceptual metaphors using various semantic memory tasks. Specifically, they examined whether the LIFE IS A JOURNEY metaphor structured participants' semantic memory of typical life events. Over a set of studies, Katz and Taylor (2008) argued that their results were suggestive of activation of the conceptual metaphor LIFE IS A JOURNEY. The supportive data included preference for forward temporal order in producing idealized life events of an imagined 70-year old, along with high agreement between participants regarding the age, the affect, and whether the event actually happened. These data suggest that knowledge of a typical life is structured as proceeding along a path with well-known landmarks along the way.

More convincing evidence can be found in episodic memory tasks. In typical episodic tasks participants are presented information to remember (study phase) and after a set amount of time are asked to remember the presented information. A well-established memory principle is that the nature of encoding the material at study will be reflected in subsequent memory. As applied to CMT, the notion is that if seemingly unrelated items activate a memory structure without conscious consideration then traces of that activation will be found in subsequent memory.

Allbritton, McKoon, and Gerrig (1995) found that recognition response times were faster in contexts that made a critical phrase metaphorical (the schema-matching condition) than the sentence as read literally, arguing that the underlying conceptual metaphor supported schematically had been automatically activated during text processing and facilitated later recognition. However these data have been criticized by McGlone (2007) who has argued that the procedure used by Allbritton et al. did not control for the possibility of strategic (not automatic) memory effects nor did the study control for a more basic explanation, namely lexical priming.

Katz and Law (2010) also found evidence for CM using the release from proactive interference (PI) procedure (Wickens, 1970). As with the typical release from PI effect (based on short-term recall of a set of words (e.g., robin, crow, sparrow), Katz and Law found that short phrases exemplifying the basic conceptual metaphor TIME IS AN EXPENDABLE RESOURCE, were recalled progressively more poorly as the instantiations of the conceptual metaphor were changed over three trials (the so-called build up of PI). When the items were changed on the $4^{th}$ trial to exemplify a different conceptual metaphor, LOVE IS A JOURNEY, a noticeable release from PI occurred. These data are completely consistent with the notion from CMT that the underlying conceptual mapping is automatically engaged when reading the metaphoric expressions. Continued exposure to the same conceptual metaphor makes that metaphor a less effective retrieval cue over trials, whereas changing conceptual metaphors (on trial 4) introduces a new retrieval cue for use. Although these findings clearly support CMT, one should nonetheless treat the positive findings cautiously as the data are based basically on only one conceptual metaphor.

In the study reported here another episodic memory task is employed, namely the DRM false memory paradigm associated with Roediger and McDermott (1995) and originally used in Deese (1959). We employ this task because it permits for a novel prediction while, at the same time affording the introduction of control conditions to discount alternative explanations. In the standard DRM task participants are presented with study lists of seemingly unrelated words, with each word on any one list strongly associated with a non-presented target word. In a later memory test of the items, the non-presented target is likely to be “falsely” remembered, that is recalled or recognized as if it had actually been on the study list. For example, one of the study lists used by Roediger and McDermott consisted of the following 15 words: bed, rest, awake, tired, dream, wake, snooze, blanket, doze, slumber, snore, nap, peace, yawn, and drowsy. The critical non-presented word for this list was “sleep,” which is associated with each word in the list. Roediger and McDermott constructed several lists similar to this and they found that participants falsely recalled the critical non-presented words at very high rates (40% and 55% for experiments 1 and 2 respectively).

There are two competing explanations for the false memory of the non-presented item. One explanation is based on spreading activation, which is the idea that when a concept is activated in semantic memory, other associated concepts are also partially activated (see Cann, McRae and Katz, 2011 for a review). In the case of the study lists used by Roediger and McDermott, each word in the list supposedly partially activates the non-presented concept, and as a result, it becomes highly active in semantic memory. When remembering the presented items at a later time the heightened level of activation for presented items, plus the non-presented target due to spreading activation, relative to non-presented items in general, is used to identify the items to be recalled or recognized.

An alternative theory for the false memory produced in the DRM paradigm follows from “fuzzy trace” theory (see Cann et al., 2011, for a review). The basic tenet of this theory is that on encountering the items in the study list two forms of representations are formed: a verbatim representation based on the details of the item and a gist representation based on the generalized (fuzzy) underlying meaning. In most cases, use of gist traces is employed in memory and reasoning tasks. In the DRM task gist recall of non-presented items occurs because the gist shares high overlap with the verbatim-based representations.

Regardless of the theoretical basis for the DRM recall of word lists, the procedure, appropriately modified, is arguably an ideal means of testing the psychological reality of conceptual metaphors. Lakoff (2008) argues that when a person encounters a metaphorical expression, the conceptual metaphor that underlies the expression is automatically activated. This assumption is similar to the notion found in experimental psychology that on encountering a word (e.g., “crow”) associative links, including superordinate category names (e.g., “BIRD”) are activated automatically. Following that logic, one can speculate that reading a metaphoric expression should automatically engage a conceptual metaphor and, if a participant reads several metaphorical expressions that, according to Lakoff and Johnson (1980), are all based on a common conceptual metaphor, this conceptual metaphor should be highly activated much like the critical non-presented words in the DRM paradigm or produce a gist representation that encompasses the meaning of the metaphoric mapping. If CMT is correct, participants should falsely remember a non-presented conceptual metaphor or other metaphoric expressions based on that conceptual metaphor after reading a list of several metaphorical expressions based on this conceptual metaphor.

In the experiment presented here, participants will read a set of lists of sentences (metaphoric expressions) with each set consisting of items that correspond to a putative conceptual metaphor identified by Lakoff and Johnson (1980). Specifically, we argue that after reading a set of phrases, such as “that cost me a day,” “how did you spend the summer break?” and “budget your hours,” CMT should predict that both the conceptual metaphor TIME IS MONEY and other non-presented instantiations of the conceptual metaphor will be falsely remembered as having occurred during the study phase at a significantly higher rate than found with falsely remembered control lures. Moreover, in line with arguments that even if the predicted results obtained they could be based not on arousal of an underlying conceptual metaphor but on associative factors, we assess also the influence of associative factors.

## 2.2 Method

### 2.2.1 Basic task

Participants were presented with several sets of sentences, each set associated with a conceptual metaphor. Shortly afterwards, for each set, participants were asked to choose the presented items from amongst a set of items, some having been presented (OLD items), some lures associated with the conceptual metaphors (the conceptual metaphor itself and other NEW items that were associated with the conceptual metaphor but had not been presented originally) and some control lures. The list structures are described in more detail below. Following the presentation and the memory testing of all the lists, a final recognition test was conducted consisting of items and lures based on each of the earlier presented lists.

### 2.2.2 Participants

Forty-eight undergraduate psychology students were recruited from Western University for participation (24 females, mean age = 18.5 years). Participants were recruited via a cloud based participant management system and participated as part of a partial course requirement. This webpage is used to connect participants with researchers -- researchers post research study advertisements and eligible participants (mostly first-year psychology students) can view descriptions of these studies, and if they so choose, sign up for available timeslots.

### 2.2.3 Materials

Five study lists were constructed based on five different well-established conceptual metaphors identified by Lakoff and Johnson (1980): IDEAS ARE FOOD, LOVE IS A JOURNEY, TIME IS MONEY, THEORIES ARE BUILDINGS, and UNDERSTANDING IS SEEING (see Appendix B for full study lists and lures). Each list had 15 phrases that were mostly taken or adjusted from the master metaphor list (Lakoff, 1994). The phrases were based on the conceptual metaphors and each phrase included one word relating to the target domain (e.g., IDEAS) and one word relating to the source domain (e.g., FOOD). In addition, none of the phrases in the study list contained either of the critical words in the conceptual metaphor, for example, none of the phrases in the IDEAS ARE FOOD list contained the words “idea” or “food.” Also, as might be obvious, the phrases were all metaphorical, so that a concept from the target domain was framed in terms of a concept from the source domain. Given that we were interested in seeing whether the conceptual metaphor itself would be falsely remembered and because these metaphors are typically written in the A IS A B form (e.g., LIFE IS A JOURNEY) whereas metaphoric expressions are rarely found in that form, we ensured that four of the fifteen phrases in each list used was presented in the nominal “a is a b” (e.g., “knowledge is consumable”). If metaphoric expressions in this form were not included in the study lists, participants could potentially use this as a strategy for correctly rejecting the critical conceptual metaphor items on the recognition test.

#### 2.2.3.1 Distractor task

As is common with this procedure, we included a distractor task in between each study list and recognition test. The distractor task consisted of 10 simple math problems that required attention to the proper order of operations to solve (e.g., $6 – 6 \div (7 – 5)$, answer = 3). Participants were asked to complete these questions mentally without using paper or a calculator. This distractor task was simply to prevent participants from rehearsing the study list phrases before memory for the study list items was assessed.

#### 2.2.3.2 Recognition tests

Following each study list and on completion of the distractor task, participants completed a recognition test on the list they had just completed. The recognition test consisted of 12 phrases of 5 different types: 5 items presented on the study list (old items), the conceptual metaphor (not presented at study), 2 new phrases consistent with the conceptual metaphor (new critical phrases not presented at study), 2 new metaphor control lures, and 2 new literal control lures. The old items were simply phrases that were presented on the preceding list, and were drawn from serial positions 1, 3, 8, 10, and 15. The conceptual metaphor items were those identified by Lakoff and Johnson (1980) that are putatively the underlying cross-domain connections that motivate all the phrases in the study list. The new critical phrases were consistent with the conceptual metaphors, but were not presented in the lists. For example, a new critical phrase for the IDEAS ARE FOOD list was "that claim is hard to swallow," which also frames an idea concept ("claim") in terms of a food concept ("swallow"), but was not one of the phrases presented on the study list. The control lures consisted of two types: metaphorical and literal. The metaphorical lures were phrases that framed the same target domain, but with a different source domain. For example, the phrase "that kind of thinking is out of style," which Lakoff categorizes under the IDEAS ARE FASHIONS conceptual metaphor, was one of the metaphorical control lures for the IDEAS ARE FOOD list. These lures were added to rule out that participants were simply encoding that the lists consisted of metaphorical phrases about a target domain (e.g., IDEAS), and not encoding the conceptual mapping of the source domain onto the target domain. For instance, if the participant simply encoded that all the phrases were metaphorical expressions about ideas, and not that the expressions specifically framed ideas in terms of food. We also included literal control lures, which were literal phrases relating to the source domain. For example, the phrase "The dessert was too sweet" is a literal statement about food, the source domain of the IDEAS ARE FOOD conceptual metaphor. These items were included to rule out that participants were simply encoding a list of sentences relating to the source domain, and not encoding the conceptual mapping (e.g., if the participant simply encoded that all the sentences were about food for the IDEAS ARE FOOD list). The critical point is that for evidence to support CMT, false recognition must involve the correct source-target mapping.

### 2.2.4 Procedure

Upon arriving at the lab, participants were given a letter of information that explained the requirements of the study, as well as a consent form to sign. The task was entirely computer based. The first two screens asked for demographic information (i.e., gender and age) and then the next three screens were instructions for the task. The participants were told that the researchers were interested in the relationship between mental math ability and memory for sentences. The purpose of this deception was so that the participants would take the distractor maths tasks seriously. The participants were also told that they would see several lists of phrases and after each list, they would have to identify items as old if it had been on the list, or new if it had not been on the list.

Following the instructions was a practice list of items to get the participants used to the task. The items on the practice list were literal phrases taken from Cardillo, Schmidt, Kranjec, and Chatterjee's (2010) matched literal and metaphor stimuli. Each phrase was presented in the centre of the screen for 3 seconds, followed by a fixation cross presented for 500 ms. The practice list was followed by the maths distractor task, which consisted of 10 short math problems. Participants were instructed to answer the questions as quickly and as accurately as possible; however, the program gave them an unlimited amount of time to answer each question. For the 10 practice math questions, participants were given feedback on what the correct answers were, but feedback was not given for subsequent trials. Following the maths test, participants completed the recognition task, in which several phrases were presented on the screen, and participants had the option of either identifying them as old by pressing the "o" key, or as new by pressing the "n" key. Similar to the math distractors, participants did not have a time constraint on the recognition test. For the practice recognition test only, participants were given feedback on whether they correctly identified each item as old or new. The five study lists followed the practice list, but the format remained the same, that is, they saw the study list, answered the maths distractors, and then completed a recognition test for the study list they had just read. Each of these three phases was presented in the same way as the practice trial, except that no feedback was given for the math distractors or the recognition tests. Following the presentation of all of the lists, there was one more math distractor consisting of 10 questions, and then there was a large recognition test at the end that tested memory for all 5 lists. The same 60 items that were already tested in the previous recognition tests were retested. In addition, 14 old items (taken from serial positions 5, 7, and 13 from each study list; however, due to an experiment error, only 2 additional old items from the TIME IS MONEY list were added from serial positions 5 and 13), 5 new consistent phrases, 5 new metaphor control lures, and 10 new literal control lures were added to this recognition test. The new lures corresponded with the five presented lists, that is, there was 1 new consistent phrase, 1 new metaphor control, and 2 new literal controls added for each study list. We could not add new conceptual metaphor lures because there is only one conceptual metaphor for each study list. The final test was to examine the longevity of the activation of the conceptual metaphor.

Following the final recognition task, participants were debriefed and thanked for their participation. The entire task took approximately 30 minutes to complete.

## 2.3 Results and discussion

### 2.3.1 Initial recognition tests

Participants completed a recognition test following each of the five study lists. One male participant identified all items as "old" and was removed from further analyses, yielding a final sample of 47 participants. The data presented below is based on performance across all study lists. That is, scores could go from 0-5 for false recognition of the conceptual metaphor, given there was only one opportunity to falsely remember that item on each of the five lists. Participants were quite accurate at correctly identifying presented items as old, the mean proportion of old items correctly categorized as old was .82[3]. Of critical interest were the proportions of falsely recognized items for the four lure types. We conducted a one-way ANOVA on these proportions. Mauchly's test of sphericity revealed that the assumption of sphericity was violated $\chi^2(5) = 34.550$, $p < .001$, thus, the degrees of freedom were adjusted using the Greenhouse-Geisser correction. The mean proportion of falsely recognized items between the lure types differed significantly, $F(2.232, 102.675) = 8.799$, $p < .001$, $\eta^2 = .161$. The mean proportion (and standard deviations) of items identified as old for each item type are displayed in Table 2.1.

[3] Due to an experiment error, two of the old items on the recognition tests varied slightly from how they were presented on the study lists. For example, on the study list one of these items was "We've come a long way as a couple," but on the recognition test it was "*They've* come a long way as a couple." We removed these items from further analyses yielding a final count of 23 old items.

**Table 2.1.** Mean proportion (and standard deviation) of items categorized as "old" across participants for each of the five item types. Means are collapsed across the five initial recognition tests.

| Lure type | Number of items in category | Mean (and SD) proportion of items categorized as old |
|---|---|---|
| Old | 23* | .8224 (.1080) |
| Critical CM | 5 | .1745 (.1939) |
| Critical consistent | 10 | .2021 (.1824) |
| Control metaphor | 10 | .1000 (.1216) |
| Control literal | 10 | .0979 (.1440) |
| **Total** | **58** | **.4101 (.0854)** |

*NOTE: As mentioned above, two old items were removed from analysis due to slight differences between study and test phases.

As can be seen, there was considerably higher proportional false memories produced for the conceptual metaphors and the new items from that for the two control lures.

We conducted four planned comparisons: critical CM's vs. metaphor controls, critical CM's vs. literal controls, critical consistent vs. metaphor controls, and critical consistent vs. literal controls. As predicted, all four comparisons were significant and in the hypothesized direction (all *p's* < 0.02).

Specifically, Least Significant Difference (LSD) post-hoc comparisons revealed that the false recognition rates did not differ significantly between the critical CM's and critical consistent phrases ($p$ = .383) or between the control metaphor lures and the control literal lures ($p$ = .875). However, there were significant differences between false recognition of the conceptual metaphor and control metaphor lures (p=.013, Cohen's *d* = .460) and control literal lures (p=.006, Cohen's *d* = .449). Similarly more false recognitions occurred for non-presented metaphor expressions that were consistent with the conceptual metaphor that both the control metaphor lures (p< .001, Cohen's *d*= .659) and control literal lures (p< .001, Cohen's *d*= .635).

### 2.3.2 Latent Semantic Analysis on lures

Recall that an expressed concern regarding CMT is that supportive results could be explained without recourse to postulating a conceptual system of cross-domain mappings. With respect to the DRM procedure in standard episodic word list studies the predominate theory postulates false memories obtain because of the high level of associative arousal. As applied to the results presented here an associative explanation could be that the critical lures (the conceptual metaphors and the new expressions that instantiate the conceptual metaphor) were falsely recognized more often than the control lures because they are semantically more related to their associated study lists. To rule this out, Latent Semantic Analysis (LSA) was used to compute the semantic distance between each lure and the sentences on its study list; LSA generates values of semantic distance between words, sentences, or texts based on the co-occurrence of words in text corpora (Landauer, Foltz, & Laham, 1998). We computed these values using the sentence comparison application on the University of Colorado's LSA website (lsa.colorado.edu) based on the most appropriate grade level, namely "General Reading up to $1^{st}$ year college" setting. This application, which computes semantic distance based on whole sentences, provided us with an LSA value indicating the semantic distance between each lure and each sentence in its respective study list (15 values in total) with higher values indicating that the sentences are closer in semantic space. We took the average of these 15 values to provide a single value for the lure that represented the semantic distance between the lure and its study list, and we did this for each of the 35 lures. The average semantic distances (and SD's) for the four lure types were as follows: critical CM = .122 (.043), critical consistent = .097 (.059), metaphor control = .081 (.080), and literal control = .094 (.058). A between-item ANOVA demonstrated that the lure types did not vary significantly in terms of average semantic distance from their respective lists, $F(3, 31) = .462$, $p = .711$.

Additional analysis demonstrated that there was not a significant correlation overall between semantic distance values and false recognition rates for the items, $r(33) = .168$, $p = .335$. Thus, we find no evidence that an associative explanation based on distance in semantic memory can account for the data we observed.

### 2.3.3 Final recognition test

Recall that participants also completed a large recognition test at the end of the study that consisted of old items and lures from all of the study lists. This recognition test consisted of the 60 items from the previous recognition tests, as well as 34 items that had not been tested yet: 14 additional old items, 5 new critical consistent lures, 5 new metaphor control lures, and 10 new literal control lures. Participants were instructed to categorize the item as "old" only if they had remembered seeing it on one of the study lists.

Participants were slightly less accurate in correctly identifying presented items as old on this final recognition test. The proportion of correctly categorized old items was .77 for items previously tested and .56 for items not previously tested. The lower accuracy for previously untested items is likely due to the greater amount of time between test and recall in which the memory traces decayed.

It was not possible to conduct a factorial ANOVA examining the effect of introducing new items because the critical conceptual metaphors had been presented as items in the original recognition tests. Because each list derives from a single conceptual metaphor, it was not possible to add new conceptual metaphors to the final recognition test, which created a missing cell in the matrix. Thus, we conducted a one-way ANOVA with seven lure types: repeated critical CM, repeated critical consistent, repeated metaphor control, repeated literal control, non-repeated critical consistent, non-repeated metaphor control, and non-repeated literal control. Mauchly's test of sphericity revealed that the assumption of sphericity was violated $\chi^2(20) = 68.423$, $p < .001$, thus, the degrees of freedom were adjusted using the Greenhouse-Geisser correction. The *F*-test revealed that the proportion of falsely recognized items varied significantly between lure types, $F(3.873, 178.150) = 40.554$, $p < .001$, $\eta^2 = .469$. The mean number and proportion (and standard deviations) of falsely recognized items for each lure type are displayed in Table 2.2. The upper panel represent items that had been employed in the original list and were repeated in the second list, whereas the lower panel represents items introduced just in the final test.

**Table 2.2.** Mean proportion (and SD's) of items categorized as old for each item type across participants in the final recognition task.

| Repeated | Item type | Number of items | Proportion categorized as old |
|---|---|---|---|
| Yes | Old | 23* | .7715 (.1452) |
| | Critical CM | 5 | .3957 (.2843) |
| | Critical consistent | 10 | .4681 (.1990) |
| | Metaphor control | 10 | .3362 (.23165) |
| | Literal control | 10 | .2106 (.2159) |
| | **Total Repeated** | **58** | **.5150 (.1223)** |
| No | Old | 14** | .5593 (.1647) |
| | Critical consistent | 5 | .2298 (.2206) |
| | Metaphor control | 5 | .0766 (.1355) |
| | Literal control | 10 | .0362 (.0705) |
| | **Total Non-repeated** | **34** | **.2860 (.0901)** |

*As mentioned above, two old items were removed from analysis due to slight differences between study and test phases.
**Due to an experiment error, only two non-repeated old items from the TIME IS MONEY list were included. Three non-repeated old items were included from the other four lists yielding a total of 14 non-repeated old items.

As is evident, items that had been previously tested showed a large increase (over 20% on average for each lure type) in false recognition on this final test. Posthoc comparisons with Bonferroni corrections were conducted on the seven categories of lures and are presented in Table 2.3.

**Table 2.3.** Posthoc comparisons with Bonferroni corrections between the seven categories of lures for the final recognition test.

| Comparison | | Mean Difference | *p*-value |
|---|---|---|---|
| Rep. Crit. CM | – Rep. Crit. consistent | -.072 | > .999 |
| | – Rep. Cont. metaphor | .060 | > .999 |
| | – Rep. Cont. literal | .185** | .004 |
| | – Non-Rep. Crit. consistent | .166* | .032 |
| | – Non-Rep. Cont. metaphor | .319*** | < .001 |
| | – Non-Rep. Cont. literal | .360*** | < .001 |
| Rep. Crit. consistent | – Rep. Cont. metaphor | .132* | .035 |
| | – Rep. Cont. literal | .257*** | < .001 |
| | – Non-Rep. Crit. consistent | .238*** | < .001 |
| | – Non-Rep. Cont. metaphor | .391*** | < .001 |
| | – Non-Rep. Cont. literal | .432*** | < .001 |
| Rep. Cont. metaphor | – Rep. Cont. literal | .126*** | < .001 |
| | – Non-Rep. Crit. consistent | .106* | .030 |
| | – Non-Rep. Cont. metaphor | .260*** | < .001 |
| | – Non-Rep. Cont. literal | .300*** | < .001 |
| Rep. Cont. literal | – Non-Rep. Crit. consistent | -.019 | > .999 |
| | – Non-Rep. Cont. metaphor | .134*** | < .001 |
| | – Non-Rep. Cont. literal | .174*** | < .001 |
| Non-Rep. Crit. consistent | – Non-Rep. Cont. metaphor | .153** | .001 |
| | – Non-Rep. Cont. literal | .194*** | < .001 |
| Non-Rep. Cont. metaphor | – Non-Rep. Cont. literal | .040 | .698 |

*$p < .05$, **$p < .01$, ***$p < .001$.

For the sake of simplicity, we will only focus on the comparisons relevant to the test of CMT. Importantly, the critical CM lures were falsely recognized at a significantly higher rate than literal controls that were repeated and all three of the non-repeated lure types. Additionally, the repeated critical consistent phrases were falsely recognized more frequently than all the other lure types (repeated and non-repeated) except for the critical conceptual metaphor lures. Also, the non-repeated critical consistent lures were falsely recognized significantly more frequently than both the non-repeated literal and metaphor controls. As with the initial recognition tests, these results support CMT as the phases consistent with the underlying conceptual mappings were falsely remembered more often than control lures.

There was one deviation from the original recognition test -- the false recognition rate for the critical CM lures ($M = .3957$, $SD = .2843$) did not differ significantly from the metaphor control lures that had been previously tested ($M = .3362$, $SD = .2316$), $p > .999$.

Recall the metaphor control lures consisted of metaphoric expressions with the same topic as the target conceptual metaphor but a different source mapping. It is not completely clear why this occurred. It may be that while the target knowledge is maintained over a longer retention period the specific mapping might be more susceptible to interference. However, this explanation could not explain the higher levels of false memories produced to new items consistent with the target conceptual metaphor compared to those produced to the metaphor controls. A more likely explanation is that the form of conceptual metaphor (A is a B) is used as a cue to reject the item as old, at least relative to items sharing the same topic.

### 2.3.4 LSA analysis on final recognition test lures

As with the initial recognition test, semantic distance of each lure to their respective study lists was computed for the final recognition test, using the same procedure as employed earlier. A one-way between-item ANOVA on the seven lure types revealed that they did not differ significantly on semantic distance to their respective lists, $F(6, 48) = 1.526$, $p = .190$. Thus, once again, the recognition memory data is not explicable by recourse to an associative explanation.

In summary, both the initial and the final recognition tests produced findings consistent with our interpretation of how CMT would be expressed in an episodic memory task. We had postulated that when one encounters a metaphorical expression (e.g., “budget your hours”), the root conceptual metaphor (TIME IS MONEY) is automatically activated, and that multiple encounters of different metaphoric expression instantiations would promote false memories both of the conceptual metaphor itself and novel instantiations. Importantly we introduced control lures in the memory tests. When the lures shared the same target domain as the conceptual metaphor but used another source domain, they were not falsely recognized as often as the critical lures. This rules out that participants were simply encoding metaphorical language about the target domain. Also, when the lures were simply literal statements about the source domain used in the conceptual metaphor (e.g., for the TIME IS MONEY list, literal statements about money), they were also not falsely recognized as often as the critical lures. This rules out that participants were simply encoding that the sentences all involved the source domain. Thus, the results suggest that both the target and source domains were encoded, even though these domains were never explicitly mentioned. Moreover, we conducted statistical analyses of semantic distance of the instances to the items on the study list, finding no such differences ruling out an explanation for the effects based on associative links.

We should mention one possibility that did not directly follow from CMT. Studies using artificial categories have shown that one can create a set of instances based on an underlying pattern, or prototype. Memory studies with these stimuli have shown that the non-presented prototype is more resistant to forgetting than learned instantiations of the prototype (Posner & Keele, 1970). If the conceptual metaphor acted similarly to a prototype then one might have expected even more false memories in the final recognition task, especially relative to the other categories of lures. However, on the final recognition test, the conceptual metaphor lures were not falsely recognized significantly more often than either the repeated critical consistent lures, or even the repeated metaphor control lures. Although the prototype studies and our task differ considerably from one another and the “A is a B” format of the conceptual metaphor differs considerably from the metaphoric expressions used with the majority of items in our task (and hence can be employed as a cue for rejecting the items) the more general question of the longevity with which a conceptual metaphor is aroused in episodic memory remains, and is a question for future research.

In summary, employing a standard episodic memory task and adopting assumptions used in the cognitive psychology literature, the data presented here provides evidence for the activation of conceptual metaphors.

Acknowledgments:

The research reported here was supported by Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant 06P0070. We thank Karen Hussey for her feedback on the research design and stimuli.

# Chapter 3

# 3 Subjective experience of the conceptual metaphor false memory effect

## 3.1 Introduction

In Chapter 2 it was demonstrated that conceptual metaphors affect episodic memory, but the study provided little insight into the nature of the false memories. For instance, it remains unclear whether participants used some type of guessing strategy to identify items as old, or whether the conceptual metaphors induced compelling false memories that felt real to the participants. The purpose of the study presented in Chapter 3 was first, to replicate the conceptual metaphor false memory effect with additional control conditions, and second, to gain further insight into the nature of this effect. The latter was accomplished by two additions. First, for items identified as old, participants were asked whether they “remember,” “know,” or “guess” the item is old (Gardiner, 1988; Tulving, 1985). Second, for half of the participants, after each list they were asked to elaborate on any strategies they used to remember the items. This was added as an indirect measure of whether participants were processing the metaphors in the lists consciously, with the logic being that if they could easily report the source-to-target metaphor mappings, they were consciously aware of the conceptual metaphors. Conversely, if participants rarely reported attending to the source-to-target mapping, but there was still evidence of a conceptual metaphor false memory effect, it would suggest that the conceptual metaphors activated automatically, without requiring conscious awareness.

### 3.1.1 The remember-know paradigm

The remember-know procedure has been used extensively in recognition memory research generally including in DRM-type tasks (Gallo, 2006). “Remember” and “know” judgments are thought to index recollection and familiarity respectively, two distinct memory processes (see Yonelinas, 2002, for a review). Recollection is when one can bring to mind details of the occurrence of an event. For a typical lab experiment, this could involve remembering what the stimulus looked like on the screen, which serial position the item was in the study list, or recalling a thought that came to mind when first seeing the item. Recollection is considered a slower process that involves intentional and conscious processes. In contrast, familiarity is when an item is confidently recognized, but the participant cannot bring to mind details of the event of experiencing the item. In contrast to recollection, familiarity is thought to be fast and automatic. Guess responses are thought to capture a weaker feeling of familiarity where the participant is not as confident in their decision (Hirshman, 1998; Yonelinas, 2002).

Remember, know, and guess judgments were employed in the current study and subsequent studies for two reasons. First, because the conceptual metaphor false memory effect is a memory phenomenon, it is important to place it in context with findings from other memory experiments, especially those on false recognition. In general, correct recognitions tend to evoke more remembering and false recognitions more knowing (Yonelinas, 2002; Brainerd & Reyna, 2005). However, in DRM tasks in which the study list items converge on a non-presented associate, participants often report "remembering" this lure. In fact, "remembering" reports are often more frequent than "knowing" reports for critical lures (Gallo, 2006; Yonelinas 2002), although the ratio of remembering to knowing is still higher for correct recognition relative to false recognition. Therefore, including remember, know, and guess judgments will provide insight into how the conceptual metaphor false memory effect relates to other false recognition effects.

Second, remember, know, and guess judgments may inform the debate on CMT and DMT, albeit indirectly. Neither theory makes explicit predictions on recollection and familiarity; however, DMT posits that cross-domain mappings depend on conscious awareness of metaphoricity. As such, one may expect that false recognition based on metaphor mappings would evoke more feelings of "remembering" as this is associated with conscious and intentional uses of memory. In contrast, CMT's position that cross-domain mappings are activated automatically and unconsciously should be associated with more "knowing" and "guessing" reports of false recognition. These responses capture feelings of familiarity, which is the faster, more automatic, and unconscious memory process. Furthermore, if conceptual metaphor activation leads to fluent processing of metaphorically consistent expressions, there should also be a larger magnitude of familiarity-based false recognitions as fluency is a heuristic for judging familiarity (Jacoby & Dallas, 1981; Jacoby & Whitehouse, 1989; Kelley & Jacoby, 1998). Consistent with this, fluency manipulations have been found to selectively influence “know” judgments (Rajaram, 1993; Rajaram & Geraci, 2000). As such, a high percentage of false recognitions attributed to “knowing” or “guessing” would suggest that conceptual metaphors are activated automatically and promote fluent processing of new expressions that use the same mappings. Note though that this is only indirect evidence and cannot carry an argument for distinguishing between CMT and DMT on its own. Nonetheless, in conjunction with the strategy descriptions, the remember, know, and guess judgments may provide some additional insight into the CMT-DMT debate.

### 3.1.2 The current study

The study lists were the same as in Chapter 2. Remember-know-guess judgments were added and half of the participants were instructed to elaborate on any strategies they used to remember the items after each recognition test. Also, rather than presenting all five study lists, only four of the five study lists were presented for each participant, and the lures associated with the non-presented study list were used as “unrelated lures” on the recognition tests. For example, this involved examining the four lure types from the TIME IS MONEY recognition test (e.g., the critical CM itself: “*time is money*”; critical consistent: “*lend me a few minutes*”; etc.), but when they are presented on the recognition test following another study list, such as IDEAS ARE FOOD. Presumably, false recognition should be much lower under these circumstances, and critically, there should be no differences in false recognition between the critical and control lures when they are all unrelated to the study list. If, for example, the critical consistent lures had a higher proportion of false recognition than the other lure types when all the lures were unrelated to the study list, it would suggest that Reid and Katz's (2018a) results were due to differences in the lures’ ability to evoke false recognition on there own and not due to these lures being consistent with the activated conceptual metaphor.

The current study also improved the design of the final recognition test by including a category of previously untested critical CM lures (i.e., the label for the conceptual metaphor itself, such as “*time is money*”). These lures supposedly capture the broad, underlying cross-domain mapping and may therefore be analogous to a prototype for the set of items (see Posner & Keele, 1970). Therefore, the memory traces for these lures may be more resistant to decay compared to the other lure types. This was not found in Reid and Katz (2018a); however, the final recognition test did not include any previously untested critical CM lures. Untested lures are the most informative as they would not have been seen until the final recognition test, and therefore, could not be contaminated from being tested previously. As such, that modification will be incorporated here.

## 3.2 Method

### 3.2.1 Participants

A total of 74 (55 female) native-English speaking participants from ages 17 to 41 (*M*age = 18.53, *SD*age = 3.39) completed the study. Participants were psychology students at Western University who completed the study as a partial course requirement. Participants were recruited through the Psychology Department's Sona system website.

### 3.2.2 Materials

#### 3.2.2.1 Study lists

The five study lists employed in Chapter 2 were used in the current study. Each participant saw only four of these five lists, yielding five different versions of the experiment based on which of the five lists was not presented. Participants were randomly assigned to the different versions. The number of participants assigned to each version are displayed in Table 3.1.

**Table 3.1.** Number of participants assigned to the different versions of the experiment. The different versions are based on the study list that was *not* presented. For instance, in the IDEAS ARE FOOD version, the IDEAS ARE FOOD study list was not presented. The lures associated with the non-presented list were used as unrelated lures on the recognition tests for the other four study lists.

| | Condition | | |
|---|---|---|---|
| Version (non-presented study list) | Strategy | No Strategy | **Total** |
| IDEAS ARE FOOD | 8 | 7 | **15** |
| LOVE IS A JOURNEY | 6 | 7 | **13** |
| THEORIES ARE BUILDINGS | 7 | 8 | **15** |
| UNDERSTANDING IS SEEING | 9* | 7 | **16** |
| TIME IS MONEY | 7 | 8 | **15** |
| **Total** | **37** | **37** | **74** |

*Note: Due to an error by the experimenter, this version had one extra participant who should have been assigned to the LOVE IS A JOURNEY version.

#### 3.2.2.2 Distractor task

The maths distractor task was the same as employed in Chapter 2.

#### 3.2.2.3 Initial recognition tests

Each recognition test consisted of eight "old" (previously presented) items and eight or nine lures. The old items were taken from serial positions 1, 3, 4, 6, 8, 10, 12, and 15 of the study list. Four types of "related" lures were employed (identical to Chapter 2) in addition to a new category of "unrelated" lures, which did not share either the target or source domain with the study list.[4] The lures associated with the recognition test of the unseen study list were used as unrelated lures on the four seen study lists. For example, if IDEAS ARE FOOD was the study list that was not presented to participants, the seven lures (one critical CM, two critical consistent, two control metaphor, and two control literal lures) that would be on this recognition test were spread across the recognition tests for the four study lists. Note that this led to an unequal division, therefore, three of the four recognition tests had two unrelated lures whereas one recognition test only had

[4] Note: For simplicity, in the results I refer to the "unrelated" lures as control lures as well, along with the two "related" control lures (control metaphor and control literal). All three of these types can be considered controls because the prediction is that the critical lures should be falsely recognized more often than all three.

one unrelated lure. Which recognition test contained only the single unrelated lure was counterbalanced across versions, such that for one version the TIME IS MONEY recognition test contained only one unrelated lure, for one version the LOVE IS A JOURNEY recognition test contained only one unrelated lure, etc.

Also, to explore whether conceptual metaphor activation continued to the final recognition test, two of the four initial recognition tests did not include the critical CM lure. This was done to allow for observations of both previously tested and untested critical CM lures on the final recognition test. Overall, each recognition test included 0 or 1 critical CM lures, 2 critical consistent lures, 2 control metaphor lures, 2 control literal lures, and 1 or 2 unrelated lures. An example of the four recognition tests for one version (non-presented study list: IDEAS ARE FOOD) of the experiment is presented in Table 3.2 below.

**Table 3.2.** Breakdown of items on the four recognition tests for one version of the experiment. IDEAS ARE FOOD was the study list not presented for this version.

| Study list | Item type | N |
|---|---|---|
| TIME IS MONEY | Old items | 8 |
| | Crit CM | 0 |
| | Crit consistent | 2 |
| | Cont metaphor | 2 |
| | Cont literal | 2 |
| | Unrelated | 2 |
| | **Total items** | **16** |
| | **Total lures** | **8** |
| LOVE IS A JOURNEY | Old items | 8 |
| | Crit CM | 1 |
| | Crit consistent | 2 |
| | Cont metaphor | 2 |
| | Cont literal | 2 |
| | Unrelated | 1 |
| | **Total items** | **16** |
| | **Total lures** | **8** |
| UNDERSTANDING IS SEEING | Old items | 8 |
| | Crit CM | 1 |
| | Crit consistent | 2 |
| | Cont metaphor | 2 |
| | Cont literal | 2 |
| | Unrelated | 2 |
| | **Total items** | **17** |
| | **Total lures** | **9** |
| THEORIES ARE BUILDINGS | Old items | 8 |
| | Crit CM | 0 |
| | Crit consistent | 2 |
| | Cont metaphor | 2 |
| | Cont literal | 2 |
| | Unrelated | 2 |
| | **Total items** | **16** |
| | **Total lures** | **8** |

#### 3.2.2.4 Final recognition test

The final recognition test included both tested and untested old items, as well as tested and untested lures for each of the five lure types. All previously tested items were retested on this final recognition test. Additionally, 20 untested old items (5 from each study list; serial positions 2, 5, 7, 13, and 14), 2 untested critical CM lures, 4 untested critical consistent lures (1 for each study list), 4 untested control metaphor lures (1 for each study list), 8 untested control literal lures (2 for each study list), and 4 untested unrelated lures were included on the final recognition test. Thus, a total of 107 items were included on the final recognition test. A breakdown of the items is displayed in Table 3.3.

**Table 3.3.** Breakdown of items on final recognition test.

| Previously tested | Item type | N |
|---|---|---|
| Yes | Old items | 32 |
| | Crit CM | 2 |
| | Crit consistent | 8 |
| | Cont metaphor | 8 |
| | Cont literal | 8 |
| | Unrelated | 7 |
| | **Total previously tested** | **65** |
| No | Old items | 20 |
| | Crit CM | 2 |
| | Crit consistent | 4 |
| | Cont metaphor | 4 |
| | Cont literal | 8 |
| | Unrelated | 4 |
| | **Total not previously tested** | **42** |

### 3.2.3 Procedure

Testing was done on Intel (processor: intel core 2 quad; screen resolution: 1440 x 900) and Asus (processor: intel core i5-7500; screen resolution: 1440 x 900) desktop computers using the E-Prime 2.0 software package. Most of the procedure was similar to Chapter 2, so I will only detail the main alterations. Along with the main task instructions, a screen was added that explained what is meant by "remember," "know," and "guess" and what type of memory would fall under each category (adjusted from Gardiner, Ramponi, & Richardson-Klavehn, 1998, see Appendix C). On the recognition tests, participants first made an old/new judgment, and if the item was identified as old, they were asked to indicate whether the item was remembered, known, or guessed (if the item was identified as new, the program simply skipped to the next item). There was no time limit for either the old/new judgments or the remember/know/guess judgments. The final recognition test followed the same procedure.

Half of the participants were asked to explain any strategies they used to remember the list. Following the recognition test for each study list, a prompt appeared on the screen that asked them to elaborate on any strategy they used, such as noticing a theme, paying attention to certain words, or repeating the items mentally (see appendix D for full instructions). The prompt also indicated that there were no right or wrong answers, and if the participant did not use a strategy, they could simple type in “I did not use any conscious strategy.” Underneath the prompt was a textbox that allowed for an open-response and there was no time limit. Only half of the participants were asked to elaborate on strategies because we did not know if consciously thinking about strategies would influence how participants performed. Therefore, for comparison the other half of participants were not asked about strategies. Participants were randomly assigned to the “strategy” and “no-strategy” conditions. The “strategy” group was only prompted after the initial recognition tests, and therefore, the final recognition test was identical for both groups.

## 3.3 Results

### 3.3.1 Initial recognition tests

#### 3.3.1.1 Preliminary analysis

I examined whether the critical CM, critical consistent, control literal, and control metaphor lures differed in false recognition when they were presented on unrelated study lists (e.g., examining a lure such as “that cost me a day,” which is associated with the TIME IS MONEY list, but when tested following another list such as IDEAS ARE FOOD).

Overall, the proportion of false recognition was infrequent when the lures were not related to the study list (critical CM: .07 critical consistent: .07, control metaphor: .08, control literal: .02; average of the four types: .05. See Appendix E for full breakdown of remember, know, and guess judgments). A 2 x 4 ANOVA was conducted with condition (strategy vs. no-strategy) as a between-subjects factor and lure type as a within-subjects factor. Critically, there was no significant main effect of lure type, $F(2.44, 175.42)$[5] =

---

[5] The assumption of sphericity was violated, therefore, the degrees of freedom were adjusted using the Greenhouse-Geisser correction.

1.59, $p = .201$. There was also no significant main effect of condition, $F(1, 72) = 0.485$, $p = .488$, nor was the interaction between condition and lure type reliable, $F(2.44, 175.42) = 2.54$, $p = .070$. These findings confirm that the four lure types did not differ significantly in terms of baseline false recognition, and therefore, if the critical lures are falsely recognized more often in the following analyses, it is attributable to conceptual metaphor activation and not to differences in how likely the lures themselves are able to elicit false recognitions.

#### 3.3.1.2 Main analysis

The main analysis is whether critical lures that further instantiate the conceptual metaphors used in the study lists (critical CM and critical consistent) are falsely recognized more often than control lures that do not instantiate the conceptual metaphor (control metaphor, control literal, unrelated). The proportions (and SD's) of items identified as old, as well as the proportions (and SD's) of remember, know, and guess judgments for the presented items and the five lure types are displayed in Table 3.4.

**Table 3.4.** Proportion (and SD) of items identified as old, as well as proportion (and SD) of items remembered, known, and guessed for initial recognition tests of Chapter 3. Note: data are collapsed across condition, as there were no significant main effects nor interactions with condition.

| Item type | Old | Remember | Know | Guess |
|---|---|---|---|---|
| Old | .86(.09) | .49(.23) | .29(.23) | .08(.08) |
| | | | | |
| Crit CM | .14(.26) | .04(.14) | .05(.16) | .05(.15) |
| Crit consistent | .28(.22) | .08(.11) | .07(.13) | .13(.16) |
| Cont metaphor | .12(.16) | .04(.09) | .03(.06) | .05(.10) |
| Cont literal | .15(.16) | .03(.07) | .04(.07) | .08(.11) |
| Unrelated | .06(.10) | .02(.05) | .01(.04) | .03(.07) |
| **Total false recognition** | **.15(.13)** | **.04(.06)** | **.04(.06)** | **.07(.07)** |

Note: Remember, know, and guess proportions are across all items tested, including items identified as new. For percentages within items identified as old, see Table 3.5.

A 2 x 5 ANOVA was conducted with lure type as a within-subjects factor and condition (strategy vs. no-strategy) as a between-subjects factor. The dependent variable was the proportion of false recognition. There was a main effect of lure type, $F(2.96, 213.29) = 21.87$, $p < .001$, $\eta^2_p = .233$, but no main effect of condition, $F(1, 72) = 0.05$, $p = .825$, nor an interaction between lure type and condition, $F(2.96, 213.29) = 1.18$, $p = .316$. Planned comparison *t*-tests (collapsed across condition) were conducted to compare the two critical lures (critical CM and critical consistent) to each of the three control lures (control metaphor, control literal, and unrelated). The alpha level was adjusted by dividing by two for these two sets of comparisons, resulting in an alpha of .025.[6] This alpha level will be used in all subsequent analyses of critical lures vs. control lures, including in the following chapters. The critical consistent lures were falsely recognized more often than all three control lures, all $t(73)$'s > 4.8, $p$'s < .001, replicating Reid and Katz (2018a). In contrast, the critical CM lures were falsely recognized more often than the unrelated lures, $t(73) = 3.00$, $p = .004$, but did not differ significantly from either the control metaphor or control literal lures, both $t(73)$'s < 0.8, $p$'s > .4.

#### 3.3.1.3 Phenomenological experience of false recognition

The percentage of remember, know, and guess judgments within correct recognitions of presented items (actual old items) and false recognitions of lures are displayed in Table 3.5.

**Table 3.5.** Percentage of items attributed remember, know, and guess judgments within correct or false recognitions. Percentages are presented by item type.

| Item type | % R | % K | % G |
|---|---|---|---|
| Old (actual presented) | 57% | 34% | 9% |
| | | | |
| Crit CM | 29% | 38% | 33% |
| Crit consistent | 28% | 26% | 47% |
| Cont metaphor | 32% | 23% | 45% |
| Cont literal | 22% | 23% | 55% |
| Unrelated | 33% | 20% | 47% |
| **Total false recognition** | **28%** | **25%** | **48%** |

A 3 (remember, know, or guess) by 5 (lure type) chi-square test was conducted to examine the percentage of remember, know, and guess judgments within the falsely

[6] For the effect to be attributed to conceptual metaphor activation the critical lure would need to be falsely recognized significantly more often than all three control lures. Therefore, the error rate is not increased by these multiple comparisons, as all three need to reach significance for the effect to be valid. For this reason, alpha was only adjusted by dividing by two for the two critical lures.

recognized lures. The chi-square test revealed that the R/K/G percentages did not vary significantly by lure type, $\chi^2(8) = 6.15$, $p = .631$. Therefore, although conceptual metaphor activation increased false recognitions for critical consistent lures, those lures did not differ reliably from the other lures with respect to the phenomenological experience of “remembering” them.

A 2 (correct vs. false recognition) x 3 (remember, know, or guess) chi-square test was also conducted to examine the phenomenological experience of correct recognition for actually presented items vs. false recognition for lures (regardless of lure type). The chi-square test was significant, $\chi^2(2) = 364.80$, $p < .001$. Posthoc tests with Bonferroni corrections (alpha = .008) revealed that all cells deviated significantly from expected values, all $z$’s > 3.3, $p$’s < .001, indicating that false recognitions were attributed significantly more as guess responses and significantly less as remember and know responses than correct recognitions.

#### 3.3.1.4 Participant strategies

There was a total of 148 strategy descriptions. Responses were categorized by two independent raters into four types depending on whether the participant reported attending to 1) just the target domain, 2) just the source domain, 3) both domains, or 4) neither domain. In their initial categorizations, the raters agreed on 88% (130 of 148) of the responses indicating high consistency. The categories of the remaining 12% (18 of 148) were decided after discussion between the two raters.

Examples of responses for each of the four categories are displayed in Table 3.6. Note that participants did not need to report the exact words used in the conceptual metaphor (e.g., LOVE or JOURNEY), but if they reported a theme closely related to one of the domains (e.g., “relationships”), it was categorized as attending to that domain.

**Table 3.6.** Examples of the four strategy types. Typos are left in. Note: not all examples were the participant's full response, but I include the relevant section for making the classification.

| Strategy type | Study list | Examples |
|---|---|---|
| Neither domain | TIME IS MONEY | "I tried to relate the sentences as I was reading them to my own life. If I link the sentences to something that I am experiencing/have experienced they should stick with me better than simply reading them." |
| | IDEAS ARE FOOD | "I repeated the sentences over and over mentally infrequently. I also looked for specific words that I thought were memorable. For example, spoon-feed, learning, eating, and raw."* |
| Target domain | TIME IS MONEY | "I noticed that many of the sentences had the common theme of time (i.e. they references hours, minutes days etc)" |
| | LOVE IS A JOURNEY | "I looked for a theme and found on in this particular list, which was about love, dating, and relationships." |
| Source domain | IDEAS ARE FOOD | "all had a reoccuring theme of food" |
| | THEORIES ARE BUILDINGS | "all of the sentences were based on a theme of construction or were about some sort of a structural element" |
| Both domains | TIME IS MONEY | "looking for a theme (in this case, time && money)" |
| | LOVE IS A JOURNEY | "All of these sentences were about relationships, so I thought about my own relationships that I have experienced while reading the sentences. I also realized that many of the sentences were using the analogy of a 'road or a car (i.i.e. bumps in the road, etc. so I was consciously not trying to be fooled by any sentences about cars when asked if I remember the new sentences." |

*Note: Although these words are related to the two domains, they were all words actually contained in the presented sentences. Because the participant did not report extracting a theme, the strategy was categorized as "neither domain."

Ninety-one of the 148 responses (61%) were categorized as attending to neither domain indicating that most of the time participants did not explicitly report consciously attending to either domain of the conceptual metaphor. The next most common response was attending to the target domain, but not mentioning the source domain (30 out of 148, or 20%). Lastly, it was fairly uncommon that the source domain (17 out of 148, or 11%) or both domains (10 out of 148, or 7%) were explicitly mentioned as a strategy.

Because it was so infrequent that participants reported attending to either the source domain or to both domains, there were not enough observations to conduct meaningful analyses on whether the strategies affected false recognition. However, I did conduct an ANOVA that excluded the trials in which the cross-domain mapping was consciously noticed, that is, when the participant reported attending to both domains. Even with these trials removed, there was still a significant main effect of lure type, $F(2.75, 98.92) = 12.30$, $p < .001$, $\eta^2_p = .255$. Planned comparison *t*-tests revealed that the proportion of false recognition for the critical consistent lures (.29) was significantly higher than for the control metaphor (.10), control literal (.16), and unrelated (.05) lures, all $t(36)$'s > 2.9, $p$'s < .01.

### 3.3.2 Final recognition test

One participant identified every item as new on the final recognition test and was therefore removed from all subsequent analyses. Thus, the analyses reported below are based on 73 participants.

#### 3.3.2.1 Preliminary analysis

The four related lure types (critical CM, critical consistent, control literal, and control metaphor) were examined to see if they differed in baseline false recognition. I first examined the previously tested lures[7] using a 2 (condition: strategy vs. no-strategy) x 4

[7] Previously tested and untested lures had to be analysed separately because there were no observations for previously untested critical CM lures unrelated to the study lists. Recall that there was only one critical CM lure for each study list, and if this study list was the one not

(lure type) ANOVA. The main effect of lure type approached significance, $F(2.70, 191.48) = 2.22$, $p = .094$, which seemed to be driven by higher false recognition for the critical consistent and control metaphor lures (.23 and .22 respectively) compared to the critical CM and control literal lures (.16 and .14 respectively; see Appendix F for full breakdown of remember, know, and guess judgments). There was no significant main effect of condition, $F(1, 71) = 0.00$, $p = .976$, nor a significant interaction between condition and lure type, $F(2.70, 191.48) = 0.58$, $p = .613$.

For the previously untested lures, there were only baseline false recognition data for three of the lure types (critical consistent: .05, control metaphor: .08, and control literal: .02). A 2 x 3 ANOVA indicated that false recognition did not vary significantly by lure type, $F(1.61, 114.06) = 1.47$, $p = .235$, nor by condition, $F(1, 71) = 1.90$, $p = .173$, nor was there a significant interaction between the two factors, $F(1.61, 114.06) = 0.20$, $p = .766$. This confirms that the lure types did not differ significantly when they were not associated to any of the presented study lists.

#### 3.3.2.2 Main analysis

The proportion (and SD) of items identified as old for actual old items as well as the five lure types for the final recognition test are displayed in Table 3.7, along with the proportion (and SD) of remember, know, and guess judgments.

---

presented to participants, then this critical CM lure would have been used as an unrelated lure on one of the initial recognition tests.

**Table 3.7.** Proportion (and SD) of items identified as old, as well as proportion (and SD) of items remembered, known, and guessed for final recognition test in Chapter 3. Proportions are presented by previous testing status (i.e., whether or not the lure had been tested in the initial recognition tests) and item type. Note: data are collapsed across condition, as there were no significant main effects nor interactions with condition.

| Previously Tested | Item type | Old | Remember | Know | Guess |
|---|---|---|---|---|---|
| Yes | Old | .80(.13) | .46(.24) | .26(.23) | .08(.09) |
| | Crit CM | .39(.33) | .17(.27) | .12(.25) | .10(.23) |
| | Crit consistent | .52(.23) | .21(.20) | .16(.18) | .15(.16) |
| | Cont metaphor | .33(.25) | .15(.20) | .10(.14) | .07(.10) |
| | Cont literal | .28(.25) | .12(.17) | .08(.16) | .08(.11) |
| | Unrelated | .19(.23) | .06(.14) | .06(.12) | .07(.11) |
| | **Total false recognition** | **.34(.19)** | **.14(.14)** | **.11(.12)** | **.09(.09)** |
| No | Old | .60(.17) | .33(.18) | .18(.16) | .09(.09) |
| | Crit CM | .18(.28) | .09(.21) | .03(.15) | .06(.17) |
| | Crit consistent | .22(.24) | .09(.16) | .05(.14) | .08(.14) |
| | Cont metaphor | .10(.16) | .04(.09) | .02(.07) | .04(.10) |
| | Cont literal | .05(.09) | .01(.04) | .01(.03) | .03(.07) |
| | Unrelated | .04(.10) | .02(.08) | .01(.04) | .02(.07) |
| | **Total false recognition** | **.12(.11)** | **.05(.08)** | **.02(.06)** | **.05(.07)** |

For overall false recognition, a 2 x 2 x 5 ANOVA was conducted with condition (strategy vs. no-strategy) as a between-subjects factor and repetition (i.e., whether the item was previously tested; yes vs. no) and lure type as within-subjects factors. There were significant main effects of both lure type, $F(2.80, 198.86) = 39.24$, $p < .001$, $\eta^2_p = .356$, and repetition, $F(1, 71) = 154.15$, $p < .001$, $\eta^2_p = .685$. There was also a significant interaction between lure type and repetition, $F(2.74, 194.48) = 2.94$, $p = .039$, $\eta^2_p = .040$. All other main effects and interactions failed to reach significance (all *F*'s < 2.7, *p*'s > .1).

The interaction between lure type and repetition indicates that the effect of repetition varied across lure type. To examine this interaction, I subtracted the proportion of false recognition for non-repeated lures from repeated lures within each of the lure types for each participant. The difference can be interpreted as the effect of repetition on false recognition; the greater the value, the greater the increase in false recognition due to repetition. I then conducted paired *t*-tests to compare the differences across the lure types. Alpha was adjusted to .005 for ten comparisons. The *t*-tests revealed that the difference was smaller for unrelated lures (.15) than for critical consistent lures (.29) and control literal lures (.23), both *t*(72)'s > 3.2, *p*'s < .003. This indicates that repetition led to a larger increase in false recognition for critical consistent and control literal lures than it did for unrelated lures. None of the other comparisons reached significance, *t*'s < 2.6, *p*'s > .01. Regarding the other two lure types, repetition led to a .21 increase in false recognition for critical CM lures and a .22 increase for control metaphor lures. Therefore, it seems that the more effective the lure is in evoking false recognition, the larger the effect of repetition. The lure with the highest proportion of false recognition (critical consistent) showed the largest increase in false recognition from repetition, and the lure with the lowest proportion of false recognition (unrelated) showed the smallest increase.

Separate sets of planned comparison *t*-tests were conducted for repeated and non-repeated lures to compare the critical lures to the three controls (alpha = .025). As there was no main effect of or interactions with condition (strategy vs. no-strategy), the data are collapsed across this variable. The critical consistent lures were falsely recognized significantly more often than all three control lures for both repeated [all *t*(72)'s > 6.3, *p*'s < .001] and non-repeated [*t*(72)'s > 3.6, *p*'s < .001] lures. For repeated lures, the critical CM lures were falsely recognized significantly more often than both the control literal and unrelated lures [both *t*(72)'s > 2.4, *p*'s < .02], but did not differ significantly from the control metaphor lures, *t*(72) = 1.61, *p* = .111. For non-repeated lures, the critical CM lures were falsely recognized significantly more often than all three control lures, all *t*(72)'s > 2.4, *p*'s < .02.

#### 3.3.2.3 Phenomenological experience of false recognition

The percentages of remember, know, and guess judgments within correctly recognized presented items and falsely recognized lures are presented in Table 3.8.

**Table 3.8.** Percentage of items attributed remember, know, and guess judgments within correct or false recognitions for the final recognition test in Chapter 3. Percentages are presented by repetition (previously tested vs. untested) and by item type.

| Previously Tested | Item Type | % R | % K | % G |
|---|---|---|---|---|
| Yes | Old (actual presented) | 57% | 33% | 10% |
| | | | | |
| | Crit CM | 44% | 32% | 25% |
| | Crit consistent | 41% | 30% | 29% |
| | Cont metaphor | 47% | 32% | 21% |
| | Cont literal | 43% | 29% | 28% |
| | Unrelated | 33% | 31% | 37% |
| | **Total false recognition** | **42%** | **31%** | **27%** |
| No | Old (actual presented) | 55% | 30% | 16% |
| | | | | |
| | Crit CM | 48% | 19% | 33% |
| | Crit consistent | 40% | 23% | 37% |
| | Cont metaphor | 37% | 20% | 43% |
| | Cont literal | 23% | 13% | 63% |
| | Unrelated | 38% | 15% | 46% |
| | **Total false recognition** | **38%** | **19%** | **43%** |

The percentage of remember, know, and guess judgments within false recognitions were examined for both repeated and non-repeated lures. Chi-square analyses (3: remember, know, or guess; by 5: lure type) revealed that the R/K/G percentages within false recognitions did not vary significantly for either repeated lures, $\chi^2(8) = 10.67$, $p = .221$, or non-repeated lures, $\chi^2(8) = 7.74$, $p = .459$. These findings parallel the results from the initial recognition tests, suggesting that once lures are falsely recognized, they produce similar phenomenological experience of recognition.

Chi-square analyses (2: correct vs. false recognition; by 3: remember, know, or guess) were also conducted to examine the phenomenological experience of correct recognition for presented items vs. false recognition for lures (regardless of lure type). The chi-square test was significant both for previously tested items, $\chi^2(2) = 146.86$, $p < .001$, and untested items, $\chi^2(2) = 66.46$, $p < .001$. Posthoc tests with Bonferroni corrections (alpha = .008) within the previously tested items indicated that the “remember” and “guess” cells for both correct and false recognition differed significantly from expected values, $z$’s > 7.3, $p$’s < .001. This indicates that false recognitions were attributed a significantly higher percentage of guess responses and significantly lower percentage of remember responses than correct recognitions. Within previously untested items, similar posthoc tests revealed that all six cells deviated significantly from expected values, $z$'s > 2.65, $p$'s < .008. This indicates that false recognitions were attributed a significantly higher percentage of guess responses and a significantly lower percentage of remember and know responses.

## 3.4 Discussion

Chapter 3 replicated the major findings described in Chapter 2. Once again, the critical consistent lures were falsely recognized significantly more often than all control lures. This finding was observed both on the initial and final recognition tests. Furthermore, this study ruled out the possibility that the conceptual metaphor false memory effect was due to differences in baseline false recognition by comparing the four related lure types when they were not associated to the presented study list.

Notably, the critical consistent lures were falsely recognized more often than controls even for non-repeated expressions on the final recognition test. The non-repeated lures are analogous to the initial recognition test lures as neither set of lures had been seen before. This essentially replicated the conceptual metaphor false memory effect for the critical consistent lures with a second set of items, indicating that the effect is fairly robust and not just due to the particular lures used in the initial recognition task. It also suggests that the effect is relatively long lasting as it occurred after the presentation of several lists.

The effect when the lure is the statement reflecting the CM itself is less robust. Unlike Chapter 2, the critical CM lures were not falsely recognized more often than control lures on the initial recognition tests. One can speculate on why these lures might induce less false recognitions than critical consistent lures. First, unlike the critical consistent lures, most of the CM lures are not conversational in nature and somewhat novel sounding. Arguably, the novelty of these expressions may be a cue to reject these items as "old." Second, these items all use the "A is B" form. Although we included a few "A is B" expressions on each study list to try to minimize this as a potential cue for rejecting critical CM lures, the majority of the study list expressions did not use this form. Therefore, this could be another feature of critical CM lures that participants use to reject these items as “old.” Finally, critical CM’s represent a superordinate category whereas the presented study list items and critical consistent lures are basic-level exemplars of the category. Lures at a different category level from study list items are known to be less effective than lures at the same category level (Park, Shobe, & Kihlstrom, 2005). For our stimuli, although some of the critical CM’s are also exemplar expressions of the mapping, such as “love is a journey” (Lakoff, 1993), most are not.[8] Therefore, the finding that the conceptual metaphors themselves are not frequently falsely remembered is actually consistent with previous research on the DRM paradigm, though it remains unclear why it occurred in the data presented in Chapter 2.

The results from the remember-know-guess judgments are both consistent and inconsistent with typical DRM studies employing word lists. False recognitions were attributed a significantly lower percentage of “remember” judgments than correct recognitions, which is typical of DRM studies (see Gallo, 2006, p. 79). However, unlike previous findings, the critical lures did not have a higher percentage of remember judgments within false recognitions than the control lures. Typically, false recognitions of critical lures have a higher percentage of remember judgments compared to false recognitions of unrelated lures in DRM studies (Gallo, 2006). Whereas some manipulations selectively influence either recollection or familiarity (e.g., Gardiner, 1988; Rajaram, 1993; Rajaram & Geraci, 2000), it appears that conceptual metaphor activation increases both illusory recollection and familiarity.

The observation here that the percentages of remember, know, and guess responses were similar across all lure types suggests that conceptual metaphor activation may be a weaker manipulation for inducing false memories that have the characteristics of “real” memories, compared to other manipulations, such as creating lists based on associative

---

[8] Lakoff (1993) points out that the conceptual metaphor labels, such as IDEAS ARE FOOD and LOVE IS A JOURNEY are “used as mnemonics to name the mappings” (p. 7). Therefore, these labels themselves are likely not part of any participant’s lexicon, except in cases like LOVE IS A JOURNEY and TIME IS MONEY in which the label of the mapping is also a conventional expression. It should be noted that the label is a name for the conceptual metaphor, which is a set of correspondences between two domains; the label itself is not the conceptual metaphor.

strength as in the classic DRM task (Roediger & McDermott, 1995; Roediger, Watson, McDermott, & Gallo, 2001), or employing lure sentences that capture the same meaning or idea as the presented sentences while also sharing many overlapping words (Bransford & Franks, 1971). Note that in the current study, the critical lures were full sentences not highly similar to their study lists in terms of word-based similarity, and that the presented items were displayed for a longer duration than is typical in DRM studies, which is a factor known to reduce false memories (Gallo & Roediger, 2002).[9] For these reasons, the current study lists and lures may not induce the compelling false recollective effects that often occur with DRM word lists.

Although the critical consistent lures did not differ from the control lures in terms of the remember-know-guess percentages, they were attributed a higher combined percentage of "know" and "guess" judgments than "remember" judgments. "Know" and "guess" judgments likely both tap into familiarity, albeit at different levels of confidence (Hirshman, 1998; Yonelinas, 2002). The high percentage of familiarity-based false recognitions are suggestive of automatic activation, which is consistent with the original CMT. This finding also supports a processing fluency explanation for the conceptual metaphor false memory effect (Jacoby & Dallas, 1981; Jacoby & Whitehouse, 1989; Kelley & Jacoby, 1998) which is consistent with automatic conceptual metaphor activation. In fact, fluency effects are eliminated when the participant is aware of the source of fluency (Jacoby & Whitehouse, 1989). That being said, familiarity-based false recognition is also consistent with fuzzy-trace theory as false alarms are assumed to be based on gist-similarity (i.e., lures being consistent with the gist, or fuzzy meaning, of the presented items), which is thought to evoke feelings of familiarity (Brainerd, Reyna, & Mojardin, 1999). Although, on face, the high percentage of familiarity-based false recognitions is consistent with the original CMT, this should not be taken as definitive evidence as the low percentage of false "remembering" could be due to the methodological reasons outlined above.

---

[9] The longer presentation duration was necessary simply because full sentences take longer to read than single words.

The strategy descriptions provide more compelling evidence for CMT over DMT as they suggest that most of the time, participants do not report attending to either domain of the conceptual metaphor. Attending to both domains was infrequently reported (7%), and even when these trials were removed, there was still evidence of a conceptual metaphor false memory effect. This suggests that the effect does not depend on the participant consciously extracting the metaphorical mapping. This finding does not support Steen's (2009) argument that cross-domain metaphor mappings only occur when participants are consciously aware a metaphor is being used, and instead aligns with the original assumption of CMT that most metaphorical cognition is unconscious (see Lakoff, 1993). Although the strategy questionnaire more strongly supports an automaticity basis for the conceptual metaphor false memory effect, the evidence is also indirect and non-experimental. Therefore, in Chapter 4, I examine automaticity more directly by manipulating attention with a concurrent task administered at study.

# Chapter 4

# 4 Examining the automaticity of the conceptual metaphor false memory effect using a divided attention task

## 4.1 Introduction

Although the findings described in the previous two chapters support CMT, it remains unclear whether conceptual metaphors are activated automatically or require conscious processing. Neither the R/K/G judgments, nor the Chapter 3 strategy analysis provided definitive answers, though the high frequency of attending to “neither” the source or target domain in the strategy analysis is more suggestive of automatic processes. To directly test the role of automatic versus conscious processing in the conceptual metaphor false memory effect, participants engaged in a concurrent task to divide attention.

Divided attention inhibits conscious and deliberate processing but typically has little to no detrimental effect on automatic processing (Jacoby, 1991; Yonelinas, 2001). Knott and Dewhurst (2007) examined the effects of divided attention on false memory with DRM lists and found that divided attention reduced false memories during the study phase, but increased them at test. Furthermore, these effects were isolated to “remember” responses; “know” responses were unaffected. Knott and Dewhurst argued that false “R” responses depend on the participants making semantic associations between the study words, and that dividing attention hinders making these associations. In contrast, dividing attention at test actually increases false “R” judgments because it inhibits controlled source monitoring decisions. Pimentel and Albuquerque (2013) divided attention during encoding of DRM lists using dichotic listening procedures and found that false memory for critical lures occurs even under minimal attention, suggesting also that the critical lures are activated automatically.

Lakoff (1993, 2008) argues that conceptual metaphors are activated automatically upon encountering metaphorical expressions. If this is the case, the expressions from the study list should activate the corresponding conceptual metaphor even under divided attention. Therefore, for Chapter 4, half of the participants engaged in a concurrent task at study that divided their attention and half completed the study with full attention. The same pattern of results for the two groups will suggest that conceptual metaphors are activated automatically, similar to the argument Pimentel and Albuquerque (2013) make for automatic activation of critical lures. However, if conceptual metaphors require deliberate and conscious processing, dividing attention should eliminate the conceptual metaphor false memory effect.

## 4.2 Method

### 4.2.1 Participants

One hundred and two participants (65 females; Sample Age: *M* = 19.87, *SD* = 4.67, range = 18-57) completed the experiment. Some participants (*N* = 64, 38 females; Sample Age: *M* = 18.45, *SD* = 0.82, range = 18-21) were recruited through the Psychology Department's Sona systems website and participated in partial fulfillment of course credit. The other participants (*N* = 38, 27 females; Sample Age: *M* = 22.26, *SD* = 7.00, range = 18-57) were recruited via posters placed around the Western University campus and were compensated $5 for participating. Preliminary analyses indicated that whether the participant was paid or not did not significantly affect overall recognition performance or the pattern of false recognition, so the recruitment type variable will not be included in any subsequent analyses. Two participants (2 females; one 18-year-old paid participant and one 19-year-old Sona participant) from the divided attention condition were removed from analysis due to poor performance on the concurrent task. These participants were replaced by two other participants yielding a final sample of 100 participants (*N* = 100, 63 females; Age: *M* = 19.90, *SD* = 4.71, range = 18-57) with 63 Sona participants (37 females; Age: *M* = 18.44, *SD* = 0.82, range = 18-21) and 37 paid participants (26 females, Age: *M* = 22.38, *SD* = 7.06, range = 18-57). The full and divided attention conditions each included 50 participants.

To ensure that the sample size was sufficiently large to detect an interactive effect of lure type and attention, a power analysis was conducted using G*Power 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007). Based on the data from the initial recognition test in Chapter 3, the correlation among repeated measures was estimated at .40 and the non-sphericity correction $\varepsilon$ was set to .741. The former value was based on the correlation of participants' false recognition proportions for the five different lure types and the latter value was based on the Greenhouse-Geisser correction. The power analysis indicated that given a total sample size of 100 and a medium effect size of $f$ = .25 (Cohen, 1988), the likelihood of detecting a significant interaction was over 99%.

### 4.2.2 Materials

The same study lists and recognition tests employed in Chapter 3 were used in the current study.

### 4.2.3 Procedure

Testing was done on an Asus (processor: intel core i5-7500; screen resolution: 1440 x 900) desktop computer using the E-Prime 2.0 software package. The study consisted of a between-subjects condition with two levels: full vs. divided attention. Due to the small number of lists in the current study, a between-subjects design was employed (as employed in Otgaar, Peters, & Howe, 2012; Pérez-Mata, Read, & Diges, 2002). Participants were randomly assigned to the full and divided attention conditions. The procedure for the full attention condition was identical to the procedure for the "no-strategy" condition described in Chapter 3. For the divided attention condition, participants engaged in a task used by Dewhurst, Barry, Swannell, Holmes, and Bathurst (2007) in a DRM task, namely to generate and say aloud a string of random numbers between 1 and 20 on pace with a metronome that ticked every 1 second at the same time they were silently reading the study list items. In this random number generating (RNG) task, participants were told to keep the numbers as random as possible with no obvious patterns or repetition and they had the opportunity to first practice this without having to read a list at the same time. The experimenter monitored the participants throughout the experiment as they did the RNG task to ensure the participants stayed on beat with the metronome and did not continually produce obvious patterns or repetitions. For both conditions, remaining aspects of the procedure were the same as employed in Chapters 2 and 3.

# 4.3 Results

## 4.3.1 Preliminary analyses

Two preliminary analyses were conducted. First, baseline false recognition for the four related lure types was examined in both conditions, as done in Chapter 3. Second, correct recognition of old items and false alarms for unrelated lures were examined between the attention conditions to ensure the divided attention task significantly influenced memory.

### 4.3.1.1 Baseline false recognition

#### 4.3.1.1.1 Initial recognition tests

The proportion of items falsely recognized for the four lure types when tested following an unrelated list are displayed in Table 4.1 (see Appendix G for proportion of items attributed remember, know, and guess judgments).

**Table 4.1.** Proportion (and SD) of false recognition by lure type when lures followed an unrelated study list on the initial recognition tests of Chapter 4.

| Condition | Lure type | Old |
|---|---|---|
| Full attention | Crit CM | .04(.20) |
| | Crit consistent | .07(.20) |
| | Cont metaphor | .05(.15) |
| | Cont literal | .05(.15) |
| | **Total full attention** | **.05(.09)** |
| Divided attention | Crit CM | .20(.40) |
| | Crit consistent | .30(.34) |
| | Cont metaphor | .31(.33) |
| | Cont literal | .24(.29) |
| | **Total divided attention** | **.26(.17)** |

Note: Totals represent the average of the four unrelated lure types' averages.

A 2 (between-subjects condition: full vs. divided attention) by 4 (within-subjects: lure type) ANOVA indicated that false recognition did not vary significantly by lure type, $F(3, 294) = 1.27$, $p = .285$. There was however a significant main effect of condition, $F(1, 98) = 58.26$, $p < .001$, $\eta^2_p = .37$, as the proportion of false recognition across all lure types was significantly higher for the divided attention group (.26) than for the full attention group (.05). However, the interaction between condition and lure type was not significant, $F(3, 294) = 0.65$, $p = .583$, indicating that there was no facilitated false recognition for any of the lure types when the conceptual metaphor was not primed at study.

#### 4.3.1.1.2 Final recognition test

The proportions of false recognition for unrelated lures are displayed in Table 4.2 (see Appendix H for full breakdown of remember, know, and guess responses).

**Table 4.2.** Proportion (and SD) of false recognition by lure type for the final recognition test when lures were unrelated to any of the presented study lists. Proportions are broken down by previous testing status and condition.

| Previously tested | Condition | Lure type | Old |
|---|---|---|---|
| Yes | Full attention | Crit CM | .12(.33) |
| | | Crit consistent | .26(.34) |
| | | Cont metaphor | .21(.32) |
| | | Cont literal | .22(.32) |
| | | **Total full attention** | **.20(.22)** |
| | Divided attention | Crit CM | .36(.48) |
| | | Crit consistent | .35(.37) |
| | | Cont metaphor | .48(.38) |
| | | Cont literal | .40(.40) |
| | | **Total divided attention** | **.40(.25)** |
| No | Full attention | Crit consistent | .04(.20) |
| | | Cont metaphor | .12(.33) |
| | | Cont literal | .02(.10) |
| | | **Total full attention** | **.06(.13)** |
| | Divided attention | Crit consistent | .12(.33) |
| | | Cont metaphor | .26(.44) |
| | | Cont literal | .16(.28) |
| | | **Total divided attention** | **.18(.24)** |

Note: Totals represent the average of the unrelated lure types' averages.

For the previously tested lures, a 2 (between-subjects condition: full vs. divided attention) by 4 (within-subjects: lure type) ANOVA revealed a main effect of condition, $F(1, 98) = 17.58$, $p < .001$, $\eta^2_p = .15$, as the divided attention group had a higher proportion of false recognitions than the full attention group (.40 vs. .20 respectively). However, once again false recognition did not vary significantly by lure type, $F(3, 294) = 1.71$, $p = .165$, and condition and lure type did not interact, $F(3, 294) = 1.41$, $p = .241$.

For the previously untested lures, a 2 by 3 (lure type; no previously untested critical CM lures) ANOVA indicated significant main effects of lure type, $F(1.69, 165.45) = 4.77$, $p = .014$, $\eta^2_p = .05$, and attention, $F(1, 98) = 9.54$, $p = .003$, $\eta^2_p = .089$, as there was a higher proportion of false recognition for the divided attention condition (.18) than the full attention condition (.06). Attention and lure type did not interact, $F(1.69, 165.45) = 0.39$, $p = .644$. Least Significant Difference posthoc tests comparing the lure types revealed that the proportion of false recognition was significantly higher for the control metaphor lures (.19) than both the critical consistent lures (.09) and control literal lures (.08), both $p$'s $< .025$. The critical consistent lures and the control literal lures did not differ significantly ($p = .770$).

In summary, the critical lures were not more likely to induce false recognitions at baseline (when unrelated to any of the presented study lists) than the control lures. Therefore, in the following analyses for both the initial and final recognition tests, if the critical lures induce more false recognitions than the control lures, it can be attributed to the presence of conceptual metaphors and cannot be attributed to pre-existing differences in the lures' abilities to evoke false recognitions, regardless of conceptual metaphor activation.

#### 4.3.1.2 Overall recognition performance

To examine the effects of the divided attention manipulation, analyses were conducted on correct recognition of presented items and false alarms to unrelated lures. Divided attention typically has a negative influence on memory, and especially on recollection (Jacoby, 1991; Yonelinas, 2001), so both recognition and remember, know, and guess judgments were examined to determine if the RNG task significantly affected memory in the standard way.

##### 4.3.1.2.1 Initial recognition tests

The proportion of items categorized as old for presented items (correct recognitions) and unrelated lures (false alarms) by condition are displayed in Table 4.3.

**Table 4.3.** Proportion of correct recognitions for presented items and false alarms for unrelated lures by condition (full vs. divided attention) for the initial recognition tests.

| Item type | Condition | Old |
|---|---|---|
| Presented | Full attention | .85(.10) |
| | Divided attention | .60(.13) |
| Unrelated[10] | Full attention | .05(.10) |
| | Divided attention | .27(.17) |

Note: Unrelated proportions are the average collapsed across lure type.

A 2 x 2 ANOVA was conducted with attention condition (full vs. divided) as a between-subjects factor, item type (presented vs. unrelated) as a within-subjects factor, and proportion identified as “old” as the dependent variable. There was a significant main effect of item type, $F(1, 98) = 868.60$, $p < .001$, $\eta^2_p = .90$, but more critically, there was a significant interaction, $F(1, 98) = 154.08$, $p < .001$, $\eta^2_p = .61$. Simple *t*-tests with Bonferroni corrections (alpha = .025) revealed that divided attention both significantly decreased correct recognitions, $t(98) = -11.13$, $p < .001$, $d = 2.23$, and increased unrelated false alarms, $t(77.30)$[11] $= 7.92$, $p < .001$, Glass’s $\Delta = 2.28$, indicating that the manipulation had the strong expected detrimental effect on memory performance.

Within correct recognitions, the phenomenological experience differed significantly between the attention conditions. A 2 (condition: divided vs. full attention) x 3 (response: remember, know, or guess) chi-square test revealed that the percentage of remember, know, and guess responses differed significantly between the two conditions (full

---

[10] Note that the unrelated proportions from Table 4.3 differ slightly from the totals in Table 4.1. This is because the totals in Table 4.1 are the average of the averages for the four unrelated subtypes. This makes the most sense for analysing differences between the unrelated types; however, because there were no main effects or interactions with lure type for unrelated lures, in the main recognition analyses the unrelated lures are collapsed across type (in other words, the average from all the unrelated lures, regardless of type). This leads to slightly different values because there are an unequal number of unrelated lure subtypes. There is only one unrelated critical CM lure, but two lures each for the other unrelated types.

[11] Degrees of freedom adjusted due to unequal variances between groups. For this reason, Glass’s Δ was used to measure effect size rather than Cohen’s *d*, the latter of which assumes equal variances between groups. Glass’s Δ uses only the standard deviation from the control group, which in this case was the full attention group.

attention: R = 54%, K = 35%, and G = 11%; divided attention: R = 47%, K = 30%, G = 23%), $\chi^2(2) = 56.03$, $p < .001$. Posthoc tests with Bonferroni corrections (alpha = .008 for examining 6 cells) revealed that the "remember" and "guess" cells in both conditions significantly deviated from the expected counts ($z$'s > 3.1, $p$'s < .008), meaning that divided attention led to an increased percentage of guesses and a decreased percentage of remember responses attributed to correctly recognized items.

#### 4.3.1.2.2 Final recognition test

The proportion of items categorized as old for presented items and unrelated lures on the final recognition test are displayed in Table 4.4.

**Table 4.4.** Proportion of correct recognition and false alarms for unrelated lures by condition (full vs. divided attention) and repetition (whether or not the item had been tested on one of the initial recognition tests).

| Item Type | Repeated | Attention | Old |
|---|---|---|---|
| Presented | Yes | Full | .77(.17) |
| | | Divided | .50(.21) |
| | No | Full | .60(.18) |
| | | Divided | .29(.18) |
| Unrelated | Yes | Full | .21(.22) |
| | | Divided | .40(.25) |
| | No | Full | .05(.11) |
| | | Divided | .18(.23) |

A 2 x 2 x 2 ANOVA was conducted with condition (full vs. divided attention) as a between-subjects factor and repetition (previously tested: yes vs. no) and item type (presented item vs. unrelated lure) as within-subjects factors. The ANOVA revealed significant main effects of condition: $F(1, 98) = 5.76$, $p = .018$, $\eta^2_p = .06$; repetition: $F(1, 98) = 158.75$, $p < .001$, $\eta^2_p = .62$; and item type, $F(1, 98) = 291.47$, $p < .001$, $\eta^2_p = .75$, and a marginally significant interaction between repetition and condition, $F(1, 98) = 3.62$, $p = .060$, $\eta^2_p = .04$. More critically however, there was a significant interaction between item type and condition, $F(1, 98) = 136.00$, $p < .001$, $\eta^2_p = .58$. Simple $t$-tests with Bonferroni corrections (alpha = .013) revealed that divided attention significantly decreased correct recognitions for presented items that were both repeated, $t(98) = -6.95$, $p < .001$, $d = 1.39$, and non-repeated, $t(98) = -8.79$, $p < .001$, $d = 1.76$, but increased false alarms for both repeated unrelated lures, $t(98) = 4.07$, $p < .001$, $d = 0.81$, and for non-repeated unrelated lures, $t(70.85) = 3.42$, $p = .001$, Glass's $\Delta = 1.11$.[12] Similar to the initial recognition tests, the decrease in correct recognition and the increase in false alarms indicates that divided attention led to significantly poorer memory. None of the other interactions reached significance (all $F$'s < 1, $p$'s > .7).

Within the correct recognitions, a 2 (condition) x 3 (response: remember, know, or guess) chi-square test revealed that the percentage of remember, know, and guess responses differed significantly between the two conditions (full attention: R = 55%, K = 34%, and G = 11%; divided attention: R = 49%, K = 27%, G = 24%), $\chi^2(2) = 84.47$, $p < .001$. Posthoc tests with Bonferroni corrections (alpha = .008) revealed that all six cells differed significantly from the expected values, all $z$'s > 2.7, $p$'s < .008. This indicates that dividing attention decreased the percentage of both remember and know judgments and increased the percentage of guess judgments. Both remember and know judgments are indicative of greater confidence in memory judgments, and therefore, the decrease in these judgments and increase in guesses suggests that even when participants were correct, they were less sure of their judgments when attention was divided.

## 4.3.2 Main analyses

### 4.3.2.1 Initial recognition tests

The proportion (and SD) of items falsely identified as old for the five lure types, and the proportion (and SD) of remember (R), know (K), and guess (G) judgments are displayed in Table 4.5.

[12] Degrees of freedom adjusted and Glass's Δ used for effect size due to unequal variances.

**Table 4.5.** Proportion of false recognition and remember, know, and guess judgments by attention condition and lure type for initial recognition tests.

| Condition | Item type | Old | R | K | G |
|---|---|---|---|---|---|
| Full attention | Crit CM | .19(.28) | .08(.19) | .06(.19) | .05(.15) |
| | Crit consistent | .28(.21) | .07(.12) | .08(.12) | .14(.16) |
| | Cont metaphor | .14(.15) | .03(.07) | .04(.09) | .07(.11) |
| | Cont literal | .11(.13) | .03(.07) | .02(.05) | .06(.09) |
| | Unrelated | .05(.10) | .01(.05) | .01(.04) | .03(.08) |
| | **Total** | **.16(.12)** | **.04(.07)** | **.04(.06)** | **.07(.07)** |
| Divided attention | Crit CM | .37(.36) | .17(.28) | .11(.21) | .09(.19) |
| | Crit consistent | .49(.21) | .18(.15) | .14(.15) | .17(.16) |
| | Cont metaphor | .37(.22) | .12(.14) | .10(.16) | .15(.16) |
| | Cont literal | .38(.23) | .14(.14) | .08(.11) | .16(.15) |
| | Unrelated | .27(.17) | .10(.14) | .06(.10) | .11(.13) |
| | **Total** | **.38(.15)** | **.14(.10)** | **.10(.10)** | **.14(.09)** |

The false recognition data here follows a trend similar to those presented in Chapters 2 and 3. For full attention, the critical consistent lures again had about twice as many false recognitions as the control metaphor and literal lures. For divided attention, the critical consistent lures were also falsely recognized more often than controls, but divided attention increased false recognitions for all lure types. Nonetheless, even with increased false recognition for controls, the proportion of false recognition for the critical consistent lures was still .12 higher than the control metaphor lures and .11 higher than the control literal lures, which is comparable to all three full attention studies (Chapter 2, 3, and the full attention condition in the current chapter).

A 2 (between-subjects: condition) x 5 (within-subjects: lure type) ANOVA with false recognition as the dependent variable revealed significant main effects of lure type, $F(2.79, 273.77) = 18.24$, $p < .001$, $\eta^2_p = .16$, and condition, $F(1, 98) = 63.49$, $p < .001$, $\eta^2_p = .39$. Critically, there was not a significant interaction, $F(2.79, 273.77) = 0.71$, $p = .539$, confirming that the pattern of false recognition across lure types did not differ appreciably in the divided and full attention conditions.

The pattern of false recognition was examined further with separate sets of planned comparisons within the two attention conditions. Alpha was divided by two for comparing the two critical lures against each of the three control lures (literal, metaphor, and unrelated), yielding an adjusted alpha of .025. The critical consistent lures were falsely recognized significantly more often than each of the three control lures for both full attention [all $t(49)$'s > 4.9, $p$'s < .001] and divided attention [all $t(49)$'s > 3.2, $p$'s < .002]. The critical CM lures were falsely recognized significantly more often than the unrelated lures under full attention, $t(49) = 3.63$, $p < .001$, but did not differ significantly from the control metaphor or control literal lures, both $t(49)$'s < 2.0, $p$'s > .06. Under divided attention, the critical CM lures did not differ significantly from any of the three control lures, all $t(49)$'s < 1.9, $p$'s > .06. The critical finding here is that under both full and divided attention, the critical consistent lures were falsely recognized more often than all three controls, suggesting that the conceptual metaphor false memory effect was still evident even when conscious processing was limited.

##### 4.3.2.1.1 Phenomenological experience of false recognition

The percentages of remember, know, and guess judgments within correctly recognized presented items and falsely recognized lures by condition are displayed in Table 4.6.

**Table 4.6.** Percentage of items attributed remember, know, and guess judgments within correct or false recognitions for the initial recognition tests in Chapter 4. Percentages are presented by attention condition (full vs. divided) and item type.

| Condition | Item type | % R | % K | % G |
|---|---|---|---|---|
| Full attention | Old (actual presented) | 54% | 35% | 11% |
| | Crit CM | 42% | 32% | 26% |
| | Crit consistent | 23% | 28% | 49% |
| | Control metaphor | 23% | 25% | 52% |
| | Control literal | 30% | 20% | 50% |
| | Unrelated | 26% | 21% | 53% |
| | **Total false recognition** | **26%** | **26%** | **48%** |
| Divided attention | Old (actual presented) | 47% | 30% | 23% |
| | Crit CM | 46% | 30% | 24% |
| | Crit consistent | 37% | 28% | 35% |
| | Control metaphor | 32% | 27% | 41% |
| | Control literal | 37% | 20% | 43% |
| | Unrelated | 36% | 22% | 42% |
| | **Total false recognition** | **36%** | **25%** | **39%** |

Separate 3 (remember, know, or guess) by 5 (lure types) chi-square tests for both attention conditions revealed that the five lure types did not differ in terms of R/K/G percentages for either full attention, $\chi^2(8) = 6.01$, $p = .646$, or divided attention, $\chi^2(8) =$

9.37, $p$ = .312. This suggests that when lures were falsely recognized, the phenomenological experience of the false recognition did not differ significantly between the different lures types. Compared to correct recognitions however, 2 (false vs. correct recognition) x 3 (remember, know, and guess) chi-square tests revealed that the R/K/G percentages differed significantly for correct recognition and false recognition under both full attention, $\chi^2(2) = 212.70$, $p < .001$, and divided attention, $\chi^2(2) = 48.21$, $p < .001$. For full attention, posthoc tests with Bonferroni corrections (alpha = .008) revealed that all six cells differed significantly from expected values, all $z$'s > 2.7, $p$'s < .008. This indicates that false recognitions were attributed a significantly higher percentage of guesses and a significantly lower percentage of both remember and know responses than correct recognitions. For divided attention, similar posthoc tests revealed that the "remember" and "guess" cells for both full attention and divided attention differed significantly from expected values, all $z$'s > 4.4, $p$'s < .001, but that the "know" cells did not, both $z$'s = 2.11, $p$'s = .035. This indicates that false recognitions were attributed a significantly higher percentage of guesses and a significantly lower percentage of remember judgments compared to correct recognitions. Therefore, relative to false recognitions and under both full attention and divided attention, correct recognitions were more based on recollection, which is consistent with previous DRM findings (Gallo, 2006).

#### 4.3.2.2 Final recognition test

The proportion of false recognition and remember, know, and guess judgments for the final recognition test are displayed in Table 4.7.

**Table 4.7.** Proportion of false recognition and remember (R), Know (K), and Guess (G) judgments, by condition, repetition (previously tested vs. untested), and lure type.

| Condition | Repeated | Item type | Old | R | K | G |
|---|---|---|---|---|---|---|
| Full attention | Yes | Crit CM | .44(.41) | .24(.34) | .11(.25) | .09(.22) |
| | | Crit consistent | .55(.26) | .20(.21) | .20(.20) | .16(.16) |
| | | Cont metaphor | .37(.22) | .15(.19) | .11(.15) | .11(.16) |
| | | Control literal | .27(.23) | .10(.15) | .10(.16) | .07(.12) |
| | | Unrelated | .21(.22) | .08(.11) | .06(.09) | .07(.12) |
| | | **Total** | **.37(.19)** | **.15(.14)** | **.11(.12)** | **.10(.09)** |
| | No | Crit CM | .21(.37) | .06(.16) | .05(.21) | .10(.25) |
| | | Crit consistent | .27(.24) | .04(.09) | .12(.17) | .12(.17) |
| | | Cont metaphor | .10(.15) | .03(.08) | .03(.08) | .05(.10) |
| | | Control literal | .05(.12) | .01(.05) | .03(.10) | .01(.03) |
| | | Unrelated | .05(.11) | .01(.04) | .02(.07) | .03(.09) |
| | | **Total** | **.14(.12)** | **.03(.05)** | **.05(.08)** | **.06(.09)** |
| Divided attention | Yes | Crit CM | .37(.40) | .18(.30) | .10(.23) | .09(.22) |
| | | Crit consistent | .47(.28) | .21(.21) | .12(.12) | .15(.17) |
| | | Cont metaphor | .44(.21) | .22(.18) | .09(.12) | .14(.15) |
| | | Control literal | .40(.23) | .18(.17) | .10(.14) | .11(.15) |
| | | Unrelated | .40(.25) | .20(.20) | .08(.14) | .12(.15) |
| | | **Total** | **.42(.21)** | **.20(.15)** | **.10(.12)** | **.12(.10)** |
| | No | Crit CM | .21(.29) | .10(.23) | .02(.10) | .09(.22) |
| | | Crit consistent | .28(.24) | .11(.17) | .11(.18) | .07(.12) |
| | | Cont metaphor | .17(.21) | .04(.09) | .03(.09) | .11(.18) |
| | | Control literal | .12(.16) | .04(.09) | .02(.05) | .07(.12) |
| | | Unrelated | .18(.23) | .05(.12) | .03(.10) | .10(.18) |
| | | **Total** | **.19(.16)** | **.07(.09)** | **.04(.07)** | **.09(.11)** |

A 2 (between-subjects condition: full vs. divided attention) x 2 (within-subjects: repeated vs. non-repeated) x 5 (within-subjects: lure type) ANOVA revealed significant main effects of both repetition, $F(1, 98) = 200.87$, $p < .001$, $\eta^2_p = .67$, and lure type, $F(2.89, 282.95) = 24.17$, $p < .001$, $\eta^2_p = .20$, as well as a significant two-way interaction between condition and lure type, $F(2.89, 282.95) = 7.62$, $p < .001$, $\eta^2_p = .07$. The main effect of condition and the other interactions did not reach significance, all *F*'s < 2.9, *p*'s > .09.

Separate sets of planned comparison *t*-tests (alpha = .025) between the critical and control lures were conducted for each of the four combinations (full attention repeated, full attention non-repeated, divided attention repeated, and divided attention non-repeated). For full attention, the critical consistent lures were falsely recognized more often than all three control lures when both repeated [all $t(49)$'s > 5.1, *p*'s < .001] and non-repeated [all $t(49)$'s > 5.3, *p*'s < .001]. The repeated critical CM lures were falsely recognized more often than the repeated control literal and unrelated lures [both $t(49)$'s $> 2.8$, $p < .01$], but did not differ significantly from the repeated control metaphor lures, $t(49) = 1.22$, $p = .229$. The non-repeated critical CM lures also were falsely recognized significantly more often than the non-repeated control literal and unrelated lures [both $t(49)$'s $> 2.8$, $p < .01$], but did not differ significantly from the non-repeated control metaphor lures, $t(49)$'s $= 2.13$, $p = .038$. In summary, under full attention the critical consistent lures were again falsely recognized more often than all three control lures, replicating the findings from Chapters 2 and 3.

In the divided attention condition, for repeated lures, the critical consistent lures were not falsely recognized significantly more often than any of the three control lures, all $t(49)$'s $< 1.9$, $p$'s $> .05$. Similarly, the repeated critical CM lures were not falsely recognized significantly more often than any of the three repeated control lures [all $t(49)$'s $< 0$][13]. For non-repeated lures however, the critical consistent lures were falsely recognized significantly more often than all three control lures [all $t(49)$'s $> 2.6$, $p$'s $< .02$]. The non-repeated critical CM lures were falsely recognized more often than the non-repeated control literal lures, $t(49) = 2.39$, $p = .021$, but did not differ significantly from either the non-repeated control metaphor or unrelated lures [both $t(49)$'s $< 1.0$, $p$'s $> .3$].

Overall, these findings replicate the previous findings from Chapters 2 and 3 that expressions that further instantiate study list conceptual metaphors ("critical consistent" lures) are falsely recognized more often than control lures. The only case in which this was not replicated was for the previously tested lures on the final recognition test in the divided attention condition, which will be considered further in the Discussion. However, in all other cases the conceptual metaphor false memory effect emerged, which highlights again the robustness of the effect, even under divided attention in which conscious processing is diminished.

---

[13] *t*-values here were actually negative as the critical CM lures were falsely recognized less often than the controls lures.

### 4.3.2.2.1 Phenomenological experience of false recognition

The percentage of remember, know, and guess judgments within correct recognitions of old items and false recognitions of lures are displayed in Table 4.8.

**Table 4.8.** Percentage of items attributed remember, know, and guess judgments within correct or false recognitions for the final recognition test in Chapter 4. Percentages are presented by repetition (previously tested vs. untested), attention conditions (full vs. divided), and item type.

| Condition | Previously tested | Item Type | % R | % K | % G |
|---|---|---|---|---|---|
| Full attention | Yes | Old (actual presented) | 57% | 34% | 10% |
| | | Crit CM | 55% | 25% | 20% |
| | | Crit consistent | 36% | 36% | 28% |
| | | Cont metaphor | 40% | 29% | 31% |
| | | Cont literal | 38% | 38% | 25% |
| | | Unrelated | 39% | 27% | 35% |
| | | **Total false recognition** | **39%** | **33%** | **28%** |
| | No | Old (actual presented) | 50% | 34% | 15% |
| | | Crit CM | 29% | 24% | 48% |
| | | Crit consistent | 15% | 43% | 43% |
| | | Cont metaphor | 30% | 25% | 45% |
| | | Cont literal | 25% | 55% | 20% |
| | | Unrelated | 10% | 40% | 50% |
| | | **Total false recognition** | **21%** | **38%** | **41%** |
| Divided attention | Yes | Old (actual presented) | 50% | 29% | 21% |
| | | Crit CM | 49% | 27% | 24% |
| | | Crit consistent | 44% | 25% | 31% |
| | | Cont metaphor | 49% | 21% | 31% |
| | | Cont literal | 46% | 26% | 28% |
| | | Unrelated | 49% | 21% | 30% |
| | | **Total false recognition** | **47%** | **23%** | **30%** |
| | No | Old (actual presented) | 46% | 21% | 33% |
| | | Crit CM | 48% | 10% | 43% |
| | | Crit consistent | 38% | 38% | 24% |
| | | Cont metaphor | 24% | 15% | 62% |
| | | Cont literal | 33% | 14% | 53% |
| | | Unrelated | 26% | 17% | 57% |
| | | **Total false recognition** | **33%** | **21%** | **46%** |

Separate 3 (remember, know, and guess) x 5 (lure type) chi-square tests were conducted on the false recognitions within each of the four combinations of attention and repetition (full-repeated, full-non-repeated, divided-repeated, divided-non-repeated). Alpha was adjusted to .013 (.05/4) for the four separate analyses. The only case in which the chi-square test reached significance was for non-repeated lures under divided attention, $\chi^2(8) = 23.58$, $p = .003$ [all other $\chi^2(8)$'s $< 9.7$, $p$'s $> .2$]. Posthoc tests with Bonferroni corrections (alpha = .003 for examining 15 cells) revealed that the "know" and "guess" cells for the critical consistent lures differed significantly from the expected values. This indicates that false recognitions of critical consistent lures were attributed a significantly higher percentage of "know" judgments and a significantly lower percentage of "guess" judgments compared to the other lure types, possibly suggesting an increased level of confidence or familiarity. However, it is unclear why this was the case only for non-repeated lures under divided attention, but none of the other combinations.

Separate 2 (false vs. correct recognition) x 3 (remember, know, or guess) chi-square tests (alpha = .013) revealed that the R/K/G percentages differed significantly between correct and false recognition in all four combinations, all $\chi^2(8)$'s $> 9.8$, $p$'s $< .01$. Posthoc tests with Bonferroni corrections (alpha = .008) revealed that in three of the combinations (full attention/repeated, full attention/non-repeated, and divided attention/non-repeated) false recognitions were attributed a significantly lower percentage of "remember" judgments and a significantly higher percentage of "guess" judgments than correct recognitions (all $z$'s $> 2.7$, $p$'s $< .008$). For repeated items under divided attention, only the "guess" cells significantly deviated from expected values (z's > 3.9, $p$'s < 001), indicating that false recognitions were attributed a significantly higher percentage of "guess" judgments compared to correct recognitions.

## 4.4 Discussion

The main finding from this study is that the critical consistent lures were falsely recognized more often than controls even under divided attention. In fact, on the initial recognition tests, there was no significant interaction between the type of lure and attention, suggesting that the pattern of false recognition for the different lures did not vary by level of attention. These data are consistent with the argument that the conceptual metaphor false memory effect does not depend on conscious, effortful processing of the sentences, at least during study. This finding converges with the finding described in Chapter 3 that participants very infrequently reported consciously attending to the metaphor mappings in their strategy descriptions. Taken together, this strongly suggests that conceptual metaphors are automatically activated, and their arousal does not require conscious and deliberate attention.

The results from the final recognition test paralleled the initial recognition test except in one case – the repeated lures under divided attention. For this group of lures, the critical consistent lures were not recognized significantly more often than control lures, thus eliminating the conceptual metaphor false memory effect. Although the lack of an effect under these circumstances was unexpected, I offer some speculation on why this was the case. The repeated lures were seen on the initial recognition tests, and therefore, although they are considered "lures" in this study, they are not truly "new" items in the sense that they were encountered previously in the experiment (albeit, not on a study list). Rejecting repeated lures is somewhat like rejecting misleading information in studies on the "misinformation effect" (Loftus, Miller, & Burns, 1978; Cann & Katz, 2005). In this paradigm, participants first witness an event (e.g., an automobile accident) and then are asked leading questions about the event that contain incorrect information. On a following memory test, participants often attribute the misleading information from the questions to the witnessed event. This can be considered a "source monitoring" error (Johnson, Hashtroudi, & Lindsay, 1993) because the information from the follow-up questions was attributed to an incorrect source (i.e., the witnessed event).

Applied to the current study, the repeated lures could be falsely recognized due to either relatedness or to failures in source monitoring, that is, misattributing the source of the lure to the study list. Divided attention has been found to reduce monitoring (Pérez-Mata et al., 2002), so it is possible that in this anomalous finding, source monitoring errors had a larger influence on false recognition and drowned out the influence of relatedness. However, within full attention, participants may have been more aware of the source of the items and better able to reject the repeated control lures. As a result, relatedness may have played a larger role in recognition decisions than when attention was divided. This may be why the conceptual metaphor false memory effect was observed only under full attention, but not under divided attention for repeated lures. Furthermore, with non-repeated lures, previously seeing the lure was not an issue as none of these lures were tested until this point. With non-repeated lures, the conceptual metaphor false memory effect was observed even under divided attention. Therefore, for this category of lures, relatedness likely also played a larger role in recognition decisions as there was not a more salient factor (i.e., having actually seen the lures) to take precedence.

Regardless of the reasons for differences between repeated and non-repeated lures, the non-repeated lures are the most informative items on the final recognition test as they were a new set of lures not yet seen, and therefore, uncontaminated by any prior testing. Under both divided and full attention, the non-repeated critical consistent lures were falsely recognized more often than the non-repeated control lures. This suggests that even under divided attention, the conceptual metaphor false memory effect is fairly long lasting, persisting after the presentation of intervening study lists.

In contrast to the critical consistent lures, the critical CM lures, that is, the labels for the non-presented conceptual metaphors themselves, were not consistently falsely recognized more often than controls. Of course, there were fewer observations of these lures per participant than the other lures as there can only be one critical CM lure per list (unlike the other lures where multiple items can be used), which makes it harder to detect significant differences in the planned comparisons. Nonetheless, in both Chapter 3 and the current study these lures were not consistently better at inducing false recognition than control lures, so it is likely they simply are not as effective as the critical consistent lures.

In terms of the phenomenological experience of false recognition, the five lure types did not consistently differ in terms of the percentage of remember, know and guess judgments. Thus, it seems that conceptual metaphor activation increases false recognition in general, but once a lure is falsely recognized, the experience is the same regardless of the type of lure it is. This is neither consistent nor inconsistent with CMT as the theory does not make specific predictions about recollection and familiarity, although the high percentage of combined “know” and “guess” judgments is suggestive of automatic activation. Unlike some manipulations that selectively affect only recollection or familiarity, it seems that conceptual metaphor activation just increases false recognition overall rather than only affecting one type of memory.

# Chapter 5

# 5 Divided attention at study and test

## 5.1 Introduction

The purpose of Chapter 5 was to further examine the automaticity of conceptual metaphor activation, particularly at retrieval. In Chapter 4, attention was divided at study, when participants were reading and encoding the sentences, but attention was full at test, when participants were retrieving the sentences from memory and making decisions whether they thought they had seen each sentence or not. Therefore, it is possible that the conceptual metaphors were consciously brought to mind at test. Perhaps conceptual metaphors are encoded unconsciously, but retrieval depends on conscious access. Furthermore, at test an actual response is required, and it is possible that the conceptual metaphors need to be consciously accessed when making a response decision to show effects on cognition.

The original CMT and DMT do not make specific predictions about encoding and retrieval. However, Glucksberg and colleagues (Glucksberg, Brown, & McGlone, 1993; Glucksberg & McGlone, 1999) have argued that in other psychological tasks, effects supportive of CMT may be due to deliberate judgments. In particular, they mention Nayak and Gibbs's (1990) idiom choice task in which participants read short paragraphs that instantiated a conceptual metaphor (e.g., ANGER IS HEAT) and then selected between two idioms to finish the paragraph, one of which was consistent with the instantiated conceptual metaphor (e.g., *"blew her top"*) and one of which was inconsistent (e.g., *"bit his head off"*; based on the ANGER IS ANIMAL BEHAVIOUR conceptual metaphor). Nayak and Gibbs found that participants more often selected the consistent idiom to finish the paragraph. This suggested that participants were sensitive to the conceptual metaphor instantiated in the preceding paragraph, providing evidence for the psychological reality of conceptual metaphors. However, Glucksberg and colleagues (Glucksberg et al., 1993; Glucksberg & McGlone, 1999) argued that the effect may depend on the participant having enough time to make a deliberate judgment about the "fit" of the idiom. In other words, conceptual metaphors may not influence online comprehension, but participants may consciously attend to the conceptual metaphors when making post-comprehension decisions and judgments about the appropriateness of idioms. Roughly mapped onto the current memory task, the implication may be that effects supportive of conceptual metaphor activation depend on conscious attention at retrieval, when recognition judgments are made.

## 5.2 Method

### 5.2.1 Participants

There were 172 participants (122 female) aged 17 to 44 ($M$ = 20.69, $SD$ = 4.27) recruited for this study. Of this number, 101 participated for $10 compensation whereas the other 71 were recruited through the Psychology Department's Sona system website and participated in partial fulfillment of course requirements. Of the 172 participants, 2 participants did not complete the study as one paid participant withdrew from the study and for one Sona participant, the fire alarm for the building went off during the study. These participants were still compensated ($10 for the paid participant and the course participation credit for the Sona participant). Two participants were removed because they completed a previous study in our lab that used similar stimuli. Two additional participants were removed due to errors by the experimenter.[14] Also, six participants were removed for poor performance on the random number generation task (two participants each from the divided attention at study, divided attention at test, and divided attention at both study and test conditions). All of these participants were replaced. However, after recruitment was completed, two additional participants (one from the divided attention at study condition and one from the full attention condition) were removed from analyses due to their performance on the recognition tests (see Results). Therefore, the final analysed sample consisted of 158 participants (112 female) aged 17 to 44 ($M$ = 20.67, $SD$ = 4.30), 93 of whom were compensated $10 and 65 of whom

[14] In one case, the experimenter did not explain the instructions for the remember-know-guess task to the participant. In the other case, the participant accidently heard the experimenter debrief the preceding participant, and therefore, knew that the study was dealing with metaphors.

completed the study for course credit. As described below, there were four independent groups and the *N*s in each group ranged from 39 to 40 (see Table 5.1 below).

I conducted a similar power analysis to that in Chapter 4, but the correlation among repeated measures was updated to .43 and the non-sphericity correction ε was updated to .720. These estimates were based on averages between Chapters 3 and 4. The power analysis indicated that given a total sample size of 156 and a medium effect size of $f$ = .25, the likelihood of detecting a significant interaction was over 99%.[15]

## 5.2.2 Materials

The same study lists and recognition items used in Chapters 3 and 4 were used in the current Chapter.

## 5.2.3 Procedure

A similar procedure as Chapter 4 was employed with the main difference being the recognition test. Rather than first making an old/new judgment and then, if the response was deemed “old,” making a remember/know/guess judgment (two-step procedure), the participant was instructed that if they thought the item was old, to press either “remember,” “know,” or “guess” directly, depending on their type of memory for the lure (one-step procedure). In other words, there was no preceding old/new judgment. If the participant thought the item was new, they were instructed to not press anything and wait for the next item to appear (as done by Knott and Dewhurst, 2007). Each recognition item was displayed for 5 seconds following a 500ms fixation cross. The time limit was imposed so that the participants under divided attention at test conditions could not compensate for the concurrent task by looking at the item for a very long time, which could allow for conscious processing. However, to maintain consistency, the one-step, timed recognition test was applied to all conditions.

---

[15] The sample size was reduced to a more conservative total of 156 because the power analysis assumes equal *N*’s in all groups.

There were two between-subjects conditions: attention at study (full vs. divided) and attention at test (full vs. divided). The resulting four combinations are shown in Table 5.1. Note that Group 1 is a replication of the studies reported in Chapters 2 and 3, whereas Group 2 is a replication of the study described in Chapter 4. The same RNG task as Chapter 4 was employed for dividing attention and participants were randomly assigned to conditions.

**Table 5.1.** Design for Chapter 5 study with divided attention at study and test. Note: Number of participants refers to the number retained in the analyses.

| | Concurrent task at Study | Concurrent task at Test | Number of Participants |
|---|---|---|---|
| Group 1 | No | No | 39 |
| Group 2 | Yes | No | 39 |
| Group 3 | No | Yes | 40 |
| Group 4 | Yes | Yes | 40 |

The final recognition test was divided into six blocks of 18 items each to allow participants a short break in between blocks. The same 107 items used in Chapters 3 and 4 were employed with the addition of one filler item (“the surprise was a hawk’s cry”) taken from Cardillo, Schmidt, Kranjec, and Chatterjee (2010) so that each block included exactly 18 items (the filler item will not be analysed). The 18 items for each block were drawn randomly from the pool of 108 items.

## 5.3 Results

### 5.3.1 Preliminary analyses

Similar preliminary analyses as conducted in Chapter 4 were conducted for baseline false recognition and correct vs. false recognition. On the initial recognition tests, one participant in the full attention condition identified every item as old and one participant from the divided attention at study condition identified all items but one (64 of 65) as old. These two participants were removed from all subsequent analyses.

#### 5.3.1.1 Baseline false recognition

##### 5.3.1.1.1 Initial recognition tests

The proportion of false recognition for the four lure types when unrelated to the study list by condition are presented in Table 5.2 (see Appendix I for full breakdown of remember, know, and guess judgments).

**Table 5.2.** Proportion of false recognition for unrelated lures by lure type and attention conditions for the initial recognition tests.

| Study condition | Test condition | Lure type | Old |
|---|---|---|---|
| Full attention | Full attention | Crit CM | .18(.39) |
| | | Crit consistent | .18(.27) |
| | | Cont metaphor | .09(.19) |
| | | Cont literal | .06(.20) |
| | | **Total** | **.13(.18)** |
| | Divided attention | Crit CM | .08(.27) |
| | | Crit consistent | .15(.26) |
| | | Cont metaphor | .10(.20) |
| | | Cont literal | .08(.21) |
| | | **Total** | **.10(.16)** |
| Divided attention | Full attention | Crit CM | .15(.37) |
| | | Crit consistent | .32(.35) |
| | | Cont metaphor | .32(.39) |
| | | Cont literal | .17(.26) |
| | | **Total** | **.24(.20)** |
| | Divided attention | Crit CM | .25(.44) |
| | | Crit consistent | .33(.33) |
| | | Cont metaphor | .35(.36) |
| | | Cont literal | .24(.34) |
| | | **Total** | **.29(.25)** |

A 2 (attention at study: full vs. divided) x 2 (attention at test: full vs. divided) x 4 (lure type) ANOVA with attention at study and attention at test as between-subjects factors and lure type as a within-subjects factor revealed a significant main effect of lure type, $F(2.77, 426.80) = 4.93$, $p = .003$, $\eta_p^2 = .03$. There was also a significant main effect of attention at study, $F(1, 154) = 22.48$, $p < .001$, $\eta_p^2 = .13$, as false recognition was higher when attention was divided at study vs. full (.27 vs. .11, respectively). The interaction between divided attention at study and lure type was not reliable, $F(2.77, 426.80) = 2.46$, $p = .067$, $\eta_p^2 = .02$, and none of the other main effects nor interactions approached significance, all $F$'s < 1.6, $p$'s > .2. Least significant difference posthoc tests comparing the lure types collapsed across attention conditions revealed that the proportion of false recognition for both the critical consistent (.24) and control metaphor (.22) lures was significantly higher than for the control literal lures (.14), both *p*’s < .01. Also, the critical consistent lures were falsely recognized more often than the critical CM lures ($p = .029$). This was likely because the critical consistent lures and control metaphor lures, like the study lists, consist of metaphorical expressions, whereas the control literal lures are literal sentences, and thus, are a different category of language from the study lists. Nonetheless, this limits the strength of any conclusions made by directly comparing these lure types in the main false recognition analyses. For this reason, comparisons will also be conducted on the adjusted false recognition scores for the critical consistent and control literal lures using the “high-threshold correction procedure,” as employed by Gallo and Roediger (2002), Gallo, Roediger, and McDermott (2001), Schacter, Verfaellie, and Pradere (1996), and Seamon, Luo, and Gallo (1998; see also Gallo, 2006, p. 31-32). The procedure is described in Appendix K.

##### 5.3.1.1.2 Final recognition test

The proportion of false recognition for the four unrelated lure types for the final recognition test is displayed in Table 5.3 (see Appendix J for a full breakdown of remember, know, and guess judgments for the unrelated lures).

**Table 5.3.** Proportion of false recognition for unrelated lures by lure type, repetition (previously tested or untested), and attention conditions for the final recognition test in Chapter 5.

| Previously tested | Attention at study | Attention at test | Lure Type | Old |
|---|---|---|---|---|
| Yes | Full | Full | Crit CM | .49(.51) |
| | | | Crit consistent | .38(.42) |
| | | | Cont metaphor | .33(.42) |
| | | | Cont literal | .36(.40) |
| | | | **Total** | **.39(.33)** |
| | | Divided | Crit CM | .38(.49) |
| | | | Crit consistent | .35(.38) |
| | | | Cont metaphor | .36(.36) |
| | | | Cont literal | .25(.36) |
| | | | **Total** | **.33(.27)** |
| | Divided | Full | Crit CM | .54(.51) |
| | | | Crit consistent | .58(.34) |
| | | | Cont metaphor | .50(.38) |
| | | | Cont literal | .33(.37) |
| | | | **Total** | **.49(.27)** |
| | | Divided | Crit CM | .53(.51) |
| | | | Crit consistent | .45(.39) |
| | | | Cont metaphor | .46(.43) |
| | | | Cont literal | .45(.39) |
| | | | **Total** | **.47(.31)** |
| No | Full | Full | Crit consistent | .05(.22) |
| | | | Cont metaphor | .10(.31) |
| | | | Cont literal | .08(.22) |
| | | | **Total** | **.08(.20)** |
| | | Divided | Crit consistent | .20(.41) |
| | | | Cont metaphor | .08(.27) |
| | | | Cont literal | .10(.26) |
| | | | **Total** | **.13(.19)** |
| | Divided | Full | Crit consistent | .15(.37) |
| | | | Cont metaphor | .28(.46) |
| | | | Cont literal | .13(.30) |
| | | | **Total** | **.19(.26)** |
| | | Divided | Crit consistent | .25(.44) |
| | | | Cont metaphor | .30(.46) |
| | | | Cont literal | .19(.33) |
| | | | **Total** | **.25(.31)** |

As there were no observations for non-repeated (not previously tested) unrelated critical CM lures, separate ANOVAs were conducted for the repeated and non-repeated lures. A 2 (attention at study: full vs. divided) x 2 (attention at test: full vs. divided) x 4 (lure type)

ANOVA conducted on the previously tested lures revealed significant main effects of both lure type, $F(2.61, 401.53) = 4.24$, $p = .008$, $\eta_p^2 = .03$, and attention at study, $F(1, 154) = 6.17$, $p = .014$, $\eta_p^2 = .04$, the latter of which was driven by a higher overall proportion of false recognition for divided attention (.48) than full attention (.36). None of the other main effects nor interactions reached significance, all *F*'s < 1.9, *p*'s > .1.

The main effect for lure type was examined further. Least significance difference posthoc tests revealed that both the critical consistent (.44) and critical CM lures (.48) had a significantly higher proportion of false recognition than the control literal lures (.35), both *p*'s < .01. As such, adjusted false recognition proportions will also be compared for these lure types in the main analyses.

A 2 (attention at study: full vs. divided) x 2 (attention at test: full vs. divided) x 3 (lure type) ANOVA on the non-repeated lures revealed a significant main effect of attention at study, $F(1, 154) = 8.78$, $p = .004$, $\eta_p^2 = .05$, as the proportion of false recognition was higher for divided attention (.22) than full attention (.10). None of the other main effects nor interactions were reliable, all *F*'s < 2.5, *p*'s > .05, and critically, there was no significant main effect of lure type, $F(1.87, 288.01) = 1.99$, $p = .142$. Therefore, the non-repeated lures can be compared directly in the main false recognition analyses without requiring adjusted proportions.

#### 5.3.1.2 Correct vs. false recognition

##### 5.3.1.2.1 Initial recognition tests

The proportion of correct recognition and false alarms for unrelated lures by attention conditions are displayed in Table 5.4.

**Table 5.4.** Proportion of items categorized as old for both presented items (correct recognitions) and unrelated lures (false alarms) by attention at study and test.

| Item Type | Study attention | Test attention | Proportion categorized as Old |
|---|---|---|---|
| Presented | Full | Full | .87(.09) |
| | | Divided | .77(.14) |
| | Divided | Full | .69(.12) |
| | | Divided | .74(.12) |
| Unrelated | Full | Full | .12(.16) |
| | | Divided | .10(.16) |
| | Divided | Full | .25(.20) |
| | | Divided | .30(.25) |

A 2 (attention at study: full vs. divided) x 2 (attention at test: full vs. divided) x 2 (item type: presented vs. unrelated lure) ANOVA revealed a significant main effect of item type, $F(1, 154) = 1052.60$, $p < .001$, $\eta_p^2 = .87$, as, unsurprisingly, presented items were more often categorized as old than unrelated lures (.77 vs. .19, respectively). More critically, there was a significant interaction between attention at study and item type, $F(1, 154) = 57.57$, $p < .001$, $\eta_p^2 = .27$, as dividing attention decreased correct recognition (.82 vs. .71) but increased false alarms (.11 vs. .28). This indicates that dividing attention at study has a strong negative influence on recognition performance. The only other interaction that reached significance was the interaction between attention at study and attention at test, $F(1, 154) = 8.25$, $p = .005$, $\eta_p^2 = .05$. This interaction suggests that when there was a match between study and test in terms of attention (e.g., full-full or divided-divided), participants were more likely to categorize items as old, regardless of item type, than when there was a mismatch (e.g., full-divided or divided-full). None of the other interactions reached significance, which suggests that unlike dividing attention at study, dividing attention at test did not have detrimental effects on recognition performance in this task.

The percentage of correct recognitions attributed remember, know, and guess judgments by attention conditions is displayed in Table 5.5.

**Table 5.5.** Percentage of items attributed remember, know, and guess responses within correctly recognized items for the initial recognition tests of Chapter 5. Percentages are presented by attention conditions at study and test.

| Attention at study | Attention at test | % R | % K | % G |
|---|---|---|---|---|
| Full | Full | 66% | 22% | 12% |
| | Divided | 77% | 15% | 8% |
| | **Total** | **71%** | **19%** | **10%** |
| Divided | Full | 39% | 32% | 29% |
| | Divided | 50% | 28% | 22% |
| | **Total** | **45%** | **30%** | **25%** |
| Total | Full | 54% | 27% | 19% |
| | Divided | 64% | 21% | 15% |
| | **Total** | **59%** | **24%** | **17%** |

Separate 2 (attention: divided vs. full) x 3 (remember, know, or guess) chi-square tests were conducted to examine the effects of dividing attention at study (collapsed across attention at test) and dividing attention at test (collapsed across attention at study). Both chi-square tests were significant, $\chi^2(2)$'s > 41, *p's* < .001, and in both cases, all six cells deviated significantly from expected values (all *z's* > 3.8, *p's* < .001). Dividing attention at study significantly decreased the percentage of remember judgments but increased the percentage of know and guess judgments. In contrast, divided attention at test significantly increased the percentage of remember judgments and decreased the percentage of know and guess judgments. These findings will be considered further in the Discussion.

#### 5.3.1.2.2 Final recognition test

The proportion of items categorized as old for presented items (correct recognition) and unrelated lures (false alarms) on the final recognition test are displayed in Table 5.6.

**Table 5.6.** Proportion of items categorized as old for both presented items (correct recognition) and unrelated lures (false alarms) for the final recognition test in Chapter 5. Proportions are presented by repetition (previously tested vs. untested), and attention conditions (full vs. divided) at study and test.

| Item type | Repeated | Study attention | Test attention | Proportion categorized as old |
|---|---|---|---|---|
| Presented | Yes | Full | Full | .81(.19) |
| | | | Divided | .76(.13) |
| | | Divided | Full | .68(.15) |
| | | | Divided | .70(.16) |
| | No | Full | Full | .65(.20) |
| | | | Divided | .64(.18) |
| | | Divided | Full | .41(.20) |
| | | | Divided | .51(.21) |
| Unrelated lure | Yes | Full | Full | .38(.33) |
| | | | Divided | .33(.26) |
| | | Divided | Full | .48(.27) |
| | | | Divided | .46(.30) |
| | No | Full | Full | .08(.20) |
| | | | Divided | .12(.19) |
| | | Divided | Full | .17(.25) |
| | | | Divided | .23(.30) |

A 2 (attention at study: full vs. divided) x 2 (attention at test: full vs. divided) x 2 (repeated: yes vs. no) x 2 (item type: presented vs. unrelated lure) ANOVA revealed a significant main effect of item type, $F(1, 154) = 434.63$, $p < .001$, $\eta_p^2 = .74$, as the presented items were more often categorized as "old" than the unrelated lures (.65 vs. .28, respectively). There was also a significant main effect of repetition, $F(1, 154) = 260.72$, $p < .001$, $\eta_p^2 = .63$, as previously tested items were more likely to be categorized as "old" than previously untested items (.57 vs. .35, respectively). There was a significant interaction between attention at study and item type, $F(1, 154) = 52.58$, $p < .001$, $\eta_p^2 = .25$. Dividing attention decreased correct recognition of presented items (full: .72; divided: .57) but increased false alarms for unrelated lures (full: .23; divided: .34). There was also a significant interaction between attention at test and repetition, $F(1, 154) = 6.72$, $p = .010$, $\eta_p^2 = .04$. Dividing attention at test decreased the proportion of repeated items categorized as old (.59 vs .56) but increased the proportion of non-repeated items categorized as old (.33 vs. .38). Finally, there was an interaction between item type and repetition, $F(1, 154) = 13.27$, $p < .001$, $\eta_p^2 = .08$. Repetition caused a greater increase in the proportion of items categorized as old for unrelated lures (.15 vs. .41; a .26 increase) than for presented items (.55 vs. .74, a .18 increase [some discrepancy due to rounding]). None of the other main effects, nor interactions reached significance, all *F*'s < 3.5, *p*'s > .05.

#### 5.3.1.2.3 Phenomenological experience of correct recognition for final recognition test

The percentage of remember, know, and guess judgments within correct recognitions by attention conditions for the final recognition test are displayed in Table 5.7.

**Table 5.7.** Percentage of items attributed remember, know, and guess responses within correctly recognized items for the final recognition test in Chapter 5. Percentages are presented by attention conditions at study and test.

| Attention at study | Attention at test | % R | % K | % G |
|---|---|---|---|---|
| Full | Full | 59% | 24% | 17% |
| | Divided | 67% | 20% | 13% |
| | **Total** | **63%** | **22%** | **15%** |
| Divided | Full | 30% | 33% | 38% |
| | Divided | 41% | 30% | 29% |
| | **Total** | **36%** | **31%** | **33%** |
| Total | Full | 46% | 28% | 26% |
| | Divided | 55% | 25% | 20% |
| | **Total** | **51%** | **26%** | **23%** |

Similar chi-square tests as conducted for the initial recognition tests revealed significant effects of attention at both study and test, both $\chi^2(2)$'s > 47, *p*'s < .001. At study, divided attention significantly decreased the percentage of remember judgments and increased the percentage of know and guess judgments (all *z*'s > 8.0, *p*'s < .001). At test, divided attention significantly increased the percentage of remember judgments and decreased the percentage of guess judgments (*z's* > 5.4, *p's* < .001).

## 5.3.2 Main analysis

### 5.3.2.1 Initial recognition tests

The proportion of false recognition, as well as proportions remembered, known, and guessed, for the five different lure types by attention conditions is displayed in Table 5.8.

**Table 5.8.** Proportion of false recognition and remember (R), Know (K), and Guess (G) judgments for the initial recognition tests in Chapter 5. Proportions are presented by attention conditions (full vs. divided) at study and test and by lure type.

| Attention at study | Attention at test | Item type | Old | R | K | G |
|---|---|---|---|---|---|---|
| Full | Full | Crit CM | .22(.34) | .06(.20) | .04(.13) | .12(.27) |
| | | Crit consistent | .29(.24) | .06(.12) | .11(.13) | .13(.15) |
| | | Cont metaphor | .18(.22) | .04(.08) | .03(.08) | .11(.15) |
| | | Cont literal | .18(.20) | .04(.09) | .03(.06) | .11(.13) |
| | | Unrelated | .12(.16) | .02(.05) | .04(.10) | .07(.10) |
| | | **Total** | **.20(.19)** | **.04(.08)** | **.05(.06)** | **.11(.12)** |
| | Divided | Crit CM | .35(.36) | .14(.28) | .10(.20) | .11(.21) |
| | | Crit consistent | .33(.23) | .14(.17) | .09(.13) | .10(.15) |
| | | Cont metaphor | .20(.20) | .09(.14) | .03(.07) | .08(.10) |
| | | Cont literal | .16(.19) | .07(.12) | .03(.07) | .07(.10) |
| | | Unrelated | .10(.16) | .06(.14) | .02(.05) | .03(.05) |
| | | **Total** | **.23(.16)** | **.10(.12)** | **.05(.07)** | **.08(.08)** |
| Divided | Full | Crit CM | .36(.36) | .05(.15) | .18(.29) | .13(.25) |
| | | Crit consistent | .54(.22) | .09(.13) | .16(.15) | .29(.16) |
| | | Cont metaphor | .38(.24) | .07(.12) | .13(.15) | .18(.16) |
| | | Cont literal | .36(.26) | .03(.06) | .17(.20) | .15(.15) |
| | | Unrelated | .25(.20) | .02(.06) | .09(.12) | .14(.13) |
| | | **Total** | **.38(.19)** | **.05(.07)** | **.15(.13)** | **.18(.11)** |
| | Divided | Crit CM | .44(.38) | .16(.26) | .16(.26) | .11(.21) |
| | | Crit consistent | .52(.26) | .17(.18) | .18(.18) | .18(.14) |
| | | Cont metaphor | .43(.25) | .14(.17) | .14(.15) | .15(.16) |
| | | Cont literal | .36(.24) | .09(.14) | .10(.13) | .17(.17) |
| | | Unrelated | .30(.25) | .04(.09) | .11(.14) | .14(.14) |
| | | **Total** | **.41(.20)** | **.12(.13)** | **.14(.12)** | **.15(.09)** |

A 2 (attention at study: full vs. divided) by 2 (attention at test: full vs. divided) x 5 (lure type) ANOVA revealed significant main effects of lure type, $F(2.76, 424.45) = 30.55$, $p < .001$, $\eta_p^2 = .17$, and attention at study, $F(1, 154) = 35.15$, $p < .001$, $\eta_p^2 = .19$, the latter of which was driven by a higher proportion of false recognition overall when, relative to full attention, attention was divided (.39 vs. .21 respectively). None of the other main effects or interactions reached significance (all *F*'s < 2.1, *p*'s > .1). Critically, the lack of significant interactions suggests that attention, either at study or test, did not influence the pattern of false recognition.

Planned *t*-test comparisons were conducted within each of the four combinations of attention (full-full, full-divided, divided-full, and divided-divided) to compare false recognition proportions for the critical lures against the control lures (alpha = .025). In all four combinations, the critical consistent lures were falsely recognized more often than the three control lures, all *t*'s > 2.3, *p*'s < .025. In contrast, the critical CM lures only differed from the three controls when attention was full at study but divided at test, all *t*(39)'s > 2.6, *p*'s < .02. In all other combinations, the critical CM lures did not differ significantly from either the control metaphor or literal lures (*t*'s < 2.3, *p*'s > .025). In summary, the conceptual metaphor false memory effect replicated under both divided and full attention conditions for the critical consistent lures.

##### 5.3.2.1.1 Adjusted comparisons

Because the critical consistent and control literal lures differed in baseline false recognition, these two lure types were compared using adjusted false recognition values (see Appendix K for more details). A 2 (attention at study: full vs. divided) x 2 (attention at test: full vs. divided) x 2 (lure type: critical consistent vs. control literal) ANOVA revealed a significant main effect of lure type, $F(1, 154) = 6.40$, $p = .012$, $\eta_p^2 = .04$, as the critical consistent lures had a higher proportion of false recognition than the control literal lures even after adjustments (.18 vs. .13 respectively). There was also a significant main effect of attention at study, $F(1, 154) = 40.24$, $p < .001$, $\eta_p^2 = .21$, as the proportion of false recognition was significantly higher under divided attention (.25) than full attention (.05). None of the other main effects nor interactions reached significance, all *F*'s < 1.2, *p*'s > .2.

Simple *t*-tests were conducted to examine whether the critical consistent lures were falsely recognized more often than the control literal lures in each of the four attention combinations after adjustments. The difference did not reach significance in any of the individual combinations, all *t*'s < 1.8, *p*'s > .05. Therefore, the difference between the critical consistent lures and control literal lures was only significant when examining across all four conditions, but the difference was not robust within any single condition on its own.

#### 5.3.2.1.2 Phenomenological experience of false recognition

The percentage of remember, know, and guess judgments within correct recognitions of actual old items and false recognitions of lures by attention conditions are displayed in Table 5.9.

**Table 5.9.** Percentage of items attributed remember, know, and guess judgments within correct or false recognitions for the initial recognition tests in Chapter 5. Percentages are presented by attention conditions (full vs. divided) at study and test and by item type.

| Attention at Study | Attention at Test | Item Type | % R | % K | % G |
|---|---|---|---|---|---|
| Full | Full | Old | 66% | 22% | 12% |
| | | Crit CM | 29% | 18% | 53% |
| | | Crit consistent | 20% | 37% | 43% |
| | | Cont metaphor | 21% | 18% | 61% |
| | | Cont literal | 25% | 18% | 58% |
| | | Unrelated | 15% | 30% | 55% |
| | | **Total false recognition** | **21%** | **26%** | **53%** |
| | Divided | Old | 77% | 15% | 8% |
| | | Crit CM | 39% | 29% | 32% |
| | | Crit consistent | 43% | 26% | 31% |
| | | Cont metaphor | 44% | 17% | 39% |
| | | Cont literal | 42% | 17% | 40% |
| | | Unrelated | 59% | 17% | 24% |
| | | **Total false recognition** | **44%** | **22%** | **34%** |
| Divided | Full | Old | 39% | 32% | 29% |
| | | Crit CM | 14% | 50% | 36% |
| | | Crit consistent | 16% | 29% | 54% |
| | | Cont metaphor | 19% | 33% | 48% |
| | | Cont literal | 9% | 48% | 43% |
| | | Unrelated | 7% | 36% | 57% |
| | | **Total false recognition** | **14%** | **37%** | **49%** |
| | Divided | Old | 50% | 28% | 22% |
| | | Crit CM | 37% | 37% | 26% |
| | | Crit consistent | 33% | 34% | 34% |
| | | Cont metaphor | 32% | 32% | 35% |
| | | Cont literal | 26% | 28% | 46% |
| | | Unrelated | 13% | 39% | 48% |
| | | **Total false recognition** | **28%** | **33%** | **38%** |

Separate 3 (remember, know, or guess) x 5 (lure type) chi-square tests were conducted within each of the attention combinations (alpha = .0125 for four tests). None of the chi-square tests reached significance [all $\chi^2(8)$'s < 20.0, *p*'s > .0125].

Separate 2 (correct vs. overall false recognition) by 3 (remember, know, or guess) chi-square tests were also conducted within the four attention combinations to compare correct and false recognition. All four tests were significant, $\chi^2(2)$'s > 75.0, *p*'s < .001. In all combinations, false recognitions were attributed a higher percentage of guess responses and a lower percentage of remember responses than correct recognitions (all *z*'s > 6.8, *p*'s < .001). Additionally, in the divided attention at test only condition, false recognitions were attributed a higher percentage of know responses than correct recognitions, but the difference was only marginally significant, *z's* = 2.64, *p's* = .008 (alpha = .008 for six cells). In the other three attention combinations, the know cells did not deviate significantly from expected values, all *z*'s < 2.2, *p*'s > .03.

#### 5.3.2.2 Final recognition test

The proportion of false recognition for the five lure types by attention conditions and repetition (previously tested vs. untested) is displayed in Table 5.10.

**Table 5.10.** Proportion of false recognition and remember (R), Know (K), and Guess (G) judgments for the final recognition test in Chapter 5. Proportions are presented by repetition (previously tested vs. untested), attention conditions (full vs. divided) at study and test, and lure type.

| Previously tested | Attention at study | Attention at test | Lure Type | Old | R | K | G |
|---|---|---|---|---|---|---|---|
| Yes | Full | Full | Crit CM | .47(.41) | .23(.38) | .14(.26) | .10(.26) |
| | | | Crit consistent | .61(.28) | .25(.29) | .16(.18) | .19(.17) |
| | | | Cont metaphor | .45(.32) | .20(.27) | .14(.15) | .11(.11) |
| | | | Cont literal | .36(.31) | .13(.23) | .11(.17) | .12(.18) |
| | | | Unrelated | .38(.33) | .13(.23) | .12(.17) | .12(.15) |
| | | | **Total** | **.45(.28)** | **.19(.25)** | **.13(.13)** | **.13(.13)** |
| | | Divided | Crit CM | .64(.34) | .30(.41) | .19(.31) | .15(.23) |
| | | | Crit consistent | .60(.25) | .28(.25) | .13(.14) | .18(.19) |
| | | | Cont metaphor | .42(.22) | .22(.20) | .10(.14) | .10(.13) |
| | | | Cont literal | .29(.26) | .14(.22) | .05(.09) | .10(.15) |
| | | | Unrelated | .33(.26) | .12(.19) | .08(.14) | .13(.15) |
| | | | **Total** | **.46(.18)** | **.21(.20)** | **.11(.10)** | **.13(.11)** |
| | Divided | Full | Crit CM | .54(.42) | .10(.23) | .23(.32) | .21(.34) |
| | | | Crit consistent | .73(.20) | .13(.18) | .31(.25) | .29(.22) |
| | | | Cont metaphor | .58(.23) | .14(.18) | .20(.18) | .24(.18) |
| | | | Cont literal | .47(.23) | .08(.14) | .16(.16) | .24(.17) |
| | | | Unrelated | .48(.27) | .11(.18) | .13(.14) | .23(.20) |
| | | | **Total** | **.56(.19)** | **.11(.14)** | **.21(.14)** | **.24(.13)** |
| | | Divided | Crit CM | .73(.28) | .24(.32) | .23(.30) | .26(.34) |
| | | | Crit consistent | .69(.23) | .25(.22) | .22(.19) | .22(.23) |
| | | | Cont metaphor | .64(.21) | .21(.20) | .22(.14) | .22(.18) |
| | | | Cont literal | .50(.27) | .13(.16) | .16(.17) | .20(.18) |
| | | | Unrelated | .46(.30) | .14(.20) | .11(.14) | .21(.21) |
| | | | **Total** | **.60(.18)** | **.19(.18)** | **.19(.11)** | **.22(.16)** |
| No | Full | Full | Crit CM | .27(.38) | .05(.15) | .04(.13) | .18(.31) |
| | | | Crit consistent | .26(.26) | .06(.13) | .06(.12) | .13(.21) |
| | | | Cont metaphor | .17(.24) | .03(.10) | .03(.08) | .12(.18) |
| | | | Cont literal | .08(.16) | .01(.05) | .00(.02) | .07(.15) |
| | | | Unrelated | .08(.20) | .03(.11) | .00(.00) | .05(.13) |
| | | | **Total** | **.17(.19)** | **.03(.09)** | **.03(.04)** | **.11(.13)** |
| | | Divided | Crit CM | .30(.35) | .15(.28) | .00(.00) | .15(.26) |
| | | | Crit consistent | .46(.29) | .18(.27) | .10(.18) | .18(.24) |
| | | | Cont metaphor | .18(.23) | .08(.20) | .03(.08) | .07(.13) |
| | | | Cont literal | .11(.19) | .07(.16) | .02(.05) | .03(.08) |
| | | | Unrelated | .12(.19) | .04(.11) | .01(.04) | .08(.14) |
| | | | **Total** | **.24(.17)** | **.10(.16)** | **.03(.05)** | **.10(.10)** |
| | Divided | Full | Crit CM | .33(.37) | .08(.24) | .14(.26) | .12(.24) |
| | | | Crit consistent | .33(.25) | .04(.10) | .12(.18) | .16(.21) |
| | | | Cont metaphor | .23(.25) | .03(.08) | .04(.11) | .17(.21) |
| | | | Cont literal | .16(.19) | .01(.03) | .04(.08) | .11(.15) |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | Unrelated | .17(.25) | .02(.07) | .03(.08) | .12(.22) |
| | | **Total** | **.24(.17)** | **.03(.07)** | **.08(.08)** | **.13(.14)** |
| | Divided | Crit CM | .44(.40) | .14(.28) | .11(.21) | .19(.27) |
| | | Crit consistent | .42(.29) | .13(.21) | .09(.14) | .21(.20) |
| | | Cont metaphor | .26(.27) | .04(.15) | .08(.18) | .13(.22) |
| | | Cont literal | .23(.25) | .05(.13) | .07(.08) | .12(.18) |
| | | Unrelated | .23(.30) | .07(.20) | .06(.14) | .11(.20) |
| | | **Total** | **.32(.23)** | **.08(.16)** | **.08(.09)** | **.15(.14)** |

A 2 (attention at study: full vs. divided) x 2 (attention at test: full vs. divided) x 2 (repeated: yes vs. no) x 5 (lure type) ANOVA revealed significant main effects of attention at study [$F(1, 154) = 12.45$, $p < .001$, $\eta_p^2 = .07$] and repetition [$F(1, 154) = 371.69$, $p < .001$, $\eta_p^2 = .71$] as false recognition was higher overall under divided attention and for repeated lures. There was also a main effect of lure type, $F(2.88, 443.97) = 72.05$, $p < .001$, $\eta_p^2 = .32$, and a significant interaction between attention at test and lure type, $F(2.88, 443.97) = 3.67$, $p = .013$, $\eta_p^2 = .02$. The interaction seemed to be due to divided attention at test increasing false recognition for the critical CM lures (.40 to .53, a .12 increase [discrepancy due to rounding]) and the critical consistent lures (.48 to .54, a .6 increase) to a greater extent than for the control metaphor (.36 to .37, a .02 increase [discrepancy due to rounding]), control literal (.27 to .28, a .01 increase) and unrelated lures (.28 to .29, a .01 increase). Simple *t*-tests comparing across test conditions (and using the false recognition proportions averaged between repeated and non-repeated items) revealed that this increase was only significant for the critical CM lures, $t(149.67)$[16] $= 2.62$, $p = .01$, but was not significant for any of the other types of lures, all *t*'s < 1.9, *p*'s > .05. There was also a significant three-way interaction between study at test, repetition, and lure type, $F(2.81, 432.82) = 4.93$, $p = .003$, $\eta_p^2 = .03$. This interaction is more difficult to interpret, but it suggests that attention at test and repetition interacted differently depending on the lure type.

Planned *t*-tests were conducted to compare each critical lure against the three control lures within each of the combinations of attention and repetition (alpha = .025). In six of

[16] Levene's test indicated unequal variances, $F = 6.67$, $p = .011$, therefore, the degrees of freedom were adjusted.

the eight combination, the critical consistent lures were falsely recognized significantly more often than all three controls (*t*'s > 3.7, *p*'s < .001). The two exceptions were that the critical consistent lures did not differ significantly from the control metaphor lures for repeated lures with divided attention at both study and test [$t(39) = 1.03$, $p = .308$] and for non-repeated lures with full attention at both study and test [$t(38) = 1.92$, $p = .062$]. Even in these cases, the differences were treading in the predicted direction and the critical consistent lures were falsely recognized significantly more often than the other two control lures (*t*'s > 4.6, *p*'s < .001). In contrast, the critical CM lures only differed significantly from all three controls in two combinations: repeated lures with divided attention at test only [*t*(39)'s > 3.9, *p*'s < .001] and non-repeated lures with divided attention at study and test [*t*(39)'s > 2.9, *p*'s < .01]. In all other combinations, the critical CM lures did not differ significantly from the control metaphor lures, *t*'s < 2.3, *p*'s > .03.

##### 5.3.2.2.1 Adjusted comparisons

A 2 (attention at study: full vs. divided) x 2 (attention at test: full vs. divided) x 2 (lure type: critical consistent vs. control literal) ANOVA was conducted on the adjusted proportions of false recognition for the repeated critical consistent and control metaphor lures (see Appendix K for more detail about the adjustment procedure employed). There was a significant main effect of attention at study, $F(1, 154) = 14.42$, $p < .001$, $\eta_p^2 = .09$, as the overall proportion of false recognition was higher under divided attention (.20) than under full attention conditions at study (.07). More critically, there was a significant main effect of lure type, $F(1, 154) = 55.31$, $p < .001$, $\eta_p^2 = .26$, as the critical consistent lures had a higher proportion of false recognition than the control literal lures even after adjustments (.22 vs. .06). None of the other main effects or interactions reached significance, all *F's* < 2.1, *p*'s > .15. In all four combinations of attention, the critical consistent lures were falsely recognized more often than the control literal lures, all *t*'s > 2.4, *p*'s < .025.

### 5.3.2.2.2 Phenomenological experience of false recognition

The percentage of remember, know, and guess judgments within correct recognitions of actual old items and false recognitions of lures for the final recognition test are displayed in Table 5.11.

**Table 5.11.** Percentage of items attributed remember, know, and guess judgments within correct or false recognitions for the final recognition test in Chapter 5. Percentages are presented by repetition (previously tested vs. untested), attention conditions (full vs. divided) at study and test, and item type.

| Previously tested | Attention at study | Attention at test | Item type | % R | % K | % G |
|---|---|---|---|---|---|---|
| Yes | Full | Full | Old (actual presented) | 64%* | 21%* | 15%* |
| | | | Crit CM | 49% | 30% | 22% |
| | | | Crit consistent | 41% | 27% | 32% |
| | | | Cont metaphor | 44% | 31% | 25% |
| | | | Cont literal | 37% | 30% | 33% |
| | | | Unrelated | 35% | 32% | 33% |
| | | | **Total false recognition** | **40%*** | **30%*** | **30%*** |
| | | Divided | Old (actual presented) | 69%* | 19% | 12%* |
| | | | Crit CM | 47% | 29% | 24% |
| | | | Crit consistent | 47% | 22% | 31% |
| | | | Cont metaphor | 53% | 23% | 24% |
| | | | Cont literal | 48% | 17% | 35% |
| | | | Unrelated | 36% | 25% | 39% |
| | | | **Total false recognition** | **47%*** | **23%** | **31%*** |
| | Divided | Full | Old (actual presented) | 32%* | 34% | 34%* |
| | | | Crit CM | 19% | 43% | 38% |
| | | | Crit consistent | 18% | 43% | 40% |
| | | | Cont metaphor | 24% | 34% | 42% |
| | | | Cont literal | 16% | 34% | 50% |
| | | | Unrelated | 24% | 27% | 49% |
| | | | **Total false recognition** | **20%*** | **36%** | **44%*** |
| | | Divided | Old (actual presented) | 44%* | 30% | 26%* |
| | | | Crit CM | 33% | 31% | 36% |
| | | | Crit consistent | 36% | 32% | 32% |
| | | | Cont metaphor | 33% | 33% | 34% |
| | | | Cont literal | 26% | 33% | 41% |
| | | | Unrelated | 31% | 24% | 45% |
| | | | **Total false recognition** | **32%*** | **31%** | **37%*** |
| No | Full | Full | Old (actual presented) | 50%* | 29%* | 22%* |

| | | | | | |
|---|---|---|---|---|---|
| | | Crit CM | 19% | 14% | 67% |
| | | Crit consistent | 23% | 25% | 53% |
| | | Cont metaphor | 15% | 15% | 70% |
| | | Cont literal | 15% | 4% | 81% |
| | | Unrelated | 33% | 0% | 67% |
| | | **Total false recognition** | **20%*** | **14%*** | **66%*** |
| | Divided | Old (actual presented) | 64%* | 23% | 13%* |
| | | Crit CM | 50% | 0% | 50% |
| | | Crit consistent | 39% | 22% | 39% |
| | | Cont metaphor | 45% | 17% | 38% |
| | | Cont literal | 61% | 14% | 25% |
| | | Unrelated | 32% | 5% | 63% |
| | | **Total false recognition** | **45%*** | **15%** | **40%*** |
| Divided | Full | Old (actual presented) | 23%* | 28% | 49% |
| | | Crit CM | 23% | 42% | 35% |
| | | Crit consistent | 14% | 37% | 49% |
| | | Cont metaphor | 11% | 17% | 72% |
| | | Cont literal | 4% | 28% | 68% |
| | | Unrelated | 11% | 19% | 70% |
| | | **Total false recognition** | **12%*** | **29%** | **59%** |
| | Divided | Old (actual presented) | 34% | 32% | 34%* |
| | | Crit CM | 31% | 26% | 43% |
| | | Crit consistent | 30% | 21% | 49% |
| | | Cont metaphor | 17% | 32% | 51% |
| | | Cont literal | 20% | 30% | 50% |
| | | Unrelated | 30% | 24% | 46% |
| | | **Total false recognition** | **25%** | **26%** | **48%*** |

*Percentage deviates significantly from expected value ($p < .008$).

Separate 3 (remember, know, or guess) x 5 (lure type) chi-square tests were conducted within each of the eight combinations of attention at study, attention at test, and repetition (alpha = .006 for eight tests). None of the chi-square tests reached significance, all $\chi^2(8) < 17.2$, *p*'s > .02.

Separate 2 (correct vs. false recognition) x 3 (remember, know, or guess) chi-square tests were also conducted within the different combinations (alpha = .006). In all eight cases, the chi-square test was significant, all $\chi^2(2)$'s > 11.0, *p*'s < .005. The cells that deviated significantly from expected values in the posthoc analyses (alpha = .008) are marked with a '*' in Table 5.14 above.

Within repeated items, a consistent pattern emerged across all combinations: false recognitions were attributed a significantly higher percentage of guess responses and a significantly lower percentage of remember responses ($z$'s > 4.1, $p$'s < .001). Within repeated items under full attention at both test and study, false recognitions were also attributed a significantly higher percentage of know responses than was observed for correct recognitions, $z$'s = 3.80, $p$'s < .001.

For non-repeated items, a different pattern emerged for each attention combination. When attention was full at study and test, false recognitions were attributed a significantly higher percentage of guess responses and a significantly lower percentage of both remember and know responses than correct recognitions, $z$'s > 3.3, $p$'s < 001. Within the divided attention at test only condition, false recognitions were attributed a significantly higher percentage of guess responses and a significantly lower percentage of remember responses than correct recognitions, $z$'s > 4.5, $p$'s < .001. Within the divided attention at study only condition, false recognitions were attributed a significantly lower percentage of remember responses than correct recognitions, $z$'s = 3.23, $p$ = .001. Finally, for divided attention at both study and test, false recognitions were attributed a significantly higher percentage of guess responses than correct recognitions, $z$'s = 3.57, $p$'s < .001.

## 5.4 Discussion

The purpose of Chapter 5 was to replicate the findings from the previous chapters and also extend the study to determine if dividing attention at test would eliminate the conceptual metaphor false memory effect. Interpretation of the results was more complicated than in the previous chapters because there were differences in the proportion of false recognition for the lure types at baseline, that is, when the lures were unrelated to the study lists. Nonetheless, even after making adjustments, across the entire study there was still evidence of a conceptual metaphor false memory effect. Importantly, the effect was not eliminated by dividing attention at either study or test. On the initial recognition tests, the critical consistent lures were falsely recognized more often than the control metaphor lures and unrelated lures in all attention conditions, and were falsely recognized more often than the control literal lures over the whole experiment, even after adjusting for the difference between these lures in baseline false recognition.

On the other hand, after adjusting for the difference in baseline false recognition between the critical consistent and control literal lures, the difference in false recognition did not reach significance in any of the four attention conditions (full-full, full-divided, divided-full, and divided-divided) on their own, but only when all four conditions were analysed together. However, each of these conditions alone was only a quarter of the data, which reduces the statistical power substantially. Critically, there were no interactions with attention at either study or test, which suggests that dividing attention did not attenuate the conceptual metaphor false memory effect. Therefore, taken together, this study provides further evidence that conceptual metaphors activate and influence memory automatically, both at encoding (study) and retrieval (test).

On the final recognition test, across all combinations of repetition and attention the critical consistent lures were falsely recognized more often than the control literal lures (even after adjustments) and the unrelated lures. In 6 of the 8 combinations, the critical consistent lures were falsely recognized more often than the control metaphor lures, and only for repeated lures with divided attention at both test and study, and for non-repeated lures with full attention at test and study did the contrasts not reach significance (and even in these cases, the differences were trending in the predicted direction). There was an interaction between attention at test and lure type on the final recognition test, due to an increase in false recognition for the critical CM lures when attention was divided. Arguably, this last finding is due to divided attention hindering participants' ability to reject these lures based on their surface form. These lures use an "A is B" format and contain fewer words on average than the study list items which may be cues to reject these items as "old". Dividing attention may hinder participants' ability to use these cues, leading to higher levels of false recognition. Importantly, there was no evidence that dividing attention at test diminished the conceptual metaphor false memory effect. Taken together, this chapter further replicates the conceptual metaphor false memory effect and provides additional evidence that the effect does not depend on conscious, strategic processing at either encoding or retrieval.

### 5.4.1 Phenomenological experience of recognition

Overall, there was no consistent pattern for the “remember, know, and guess” percentages within false recognitions for the different lure types. Therefore, it seems that when the lures actually are falsely recognized, there is no special status for the critical consistent lures in the R/K/G percentages. For instance, these lures do not have a higher percentage of recollection (R) or familiarity (K) than the falsely recognized lures from the other lure types. The R/K/G judgments were included as exploratory since CMT makes no predictions on whether conceptual metaphor activation should primarily influence recollection or familiarity. That being said, the critical consistent lures always have less “remember” judgments than they do combined “know” and “guess” judgments. This suggests that most of the time the lures induce feelings of familiarity, rather than a quasi-sensory experience of recollection. This finding is consistent with both processing fluency (Jacoby & Whitehouse, 1989) and fuzzy-trace theory explanations (Brainerd & Reyna, 2005) of false memory as both theories suggest that false memories are based primarily on familiarity. Furthermore, familiarity is thought to be more automatic whereas recollection is thought to involve more effortful processing (Yonelinas, 2002). Therefore, the high rates of familiarity indicated by “know” and “guess” responses further support the automatic activation of conceptual metaphors, albeit indirectly.

One additional finding for correctly recognized items was that the percentage of “remember” judgments decreased when attention was divided at study but increased when attention was divided at test. Although seemingly contradictory, Knott and Dewhurst (2007) obtained similar findings. They suggest that at test, “remember” judgments can be made rapidly whereas “know” judgments require postretrieval decisions that are disrupted by divided attention. Arguably, the same might be happening with the processing of the metaphor lists.

# Chapter 6

## 6 General discussion

The purpose of this thesis was to empirically investigate the psychological reality of conceptual metaphors by testing whether reading a list of expressions based on a presumed conceptual metaphor would activate other non-presented expressions of that metaphor. Across all four experiments, it was found that this was the case: participants consistently falsely recognized metaphorically related expressions more often than topically related control expressions. The control lures were related in terms of either the target domain (control metaphor) or source domain (control literal) and were about equally related to the study list expressions in terms of word-based similarity. The findings are thus completely consistent with the argument that the conceptual metaphor false memory effect depends upon the expression using the same source-to-target metaphor mapping. More generally, the findings support the claim (Gibbs, 1996, 2011, 2013) that conceptual metaphors are indeed psychologically real and influence cognition.

The secondary purpose was to examine another controversial question: Do conceptual metaphors influence memory automatically and unconsciously? This was examined here using both direct and indirect measures and converging evidence indicates that conceptual metaphor activation is automatic. In Chapter 3, although participants rarely reported consciously attending to source-to-target metaphor mappings, the results from the recognition tests clearly replicated the conceptual metaphor false memory effect found in Reid & Katz (2018a). Furthermore, in Chapters 4 and 5, dividing attention did not diminish the conceptual metaphor false memory effect. These results taken together strongly suggest that conceptual metaphor activation is automatic and unconscious. This is consistent with the original CMT (Lakoff & Johnson, 1980; Lakoff, 1993) and is also consistent with the finding from Katz and Law (2010) that giving participants hints about the metaphorical nature of the to-be-remembered expressions (and thus, bringing conscious attention to the metaphor mappings) did not increase memory performance, but in fact, hindered performance.

In general, the findings from the current set of studies align with previous research that strongly suggests the traditional word-based DRM false memory effect is automatic. Along with findings that the effect still occurs under divided attention (Knott & Dewhurst, 2007; Pimentel & Albuquerque, 2013), it also occurs when participants are warned that the purpose of the study is to induce a false memory (Gallo, Roberts, & Seamon, 1997; Gallo, Roediger, & McDermott, 2001; McDermott & Roediger, 1998; Multhaup & Conner, 2002) and when study list words are presented rapidly (20ms; Seamon, Luo, & Gallo, 1998; Kawasaki & Yama, 2006) such that the participant is not consciously aware of the words being presented. These findings suggest that false memories for critical lures in the DRM paradigm occur automatically and unconsciously, and that participants have difficulty resisting the memory illusion even when they are expecting it, at least with word list stimuli. The findings presented in this thesis indicating that the conceptual metaphor false memory effect occurs automatically and not consciously thus align nicely with the body of research examining the DRM effect.

## 6.1 Implications for DRM research

Most of the previous research on the DRM paradigm has focused on lists of words. The current study extends the DRM paradigm to full sentences, and furthermore, complex sentences involving figurative language. An analogous memory illusion was found by Bransford and Franks (1971) in which non-presented sentences were falsely recognized if they contained highly similar semantic content to the presented sentence. In that study, however, the sentence lures were much more obviously related and shared many of the same words. In the current study, the sentence lures were less related, not sharing as much overlap in terms of words, and the sentences were not all referring to the same event or occurrence. However, the sentences were related in terms of a deeper underlying meaning, in this case sharing a common cross-domain metaphorical mapping. Therefore, this finding extends previous research on false recognition for sentences by suggesting that even when sentences are fairly unrelated on the surface, they can still induce false recognition if they are based on the same metaphorical mapping.

Although the goal of this thesis was not to test alternative false memory theories, I speculate on how three major theories, automatic spreading activation, processing fluency, and fuzzy-trace theory, could accommodate the findings. Following this, I consider the findings from the “remember,” “know,” and “guess” judgments and how they may relate to CMT and DMT.

### 6.1.1 Automatic Spreading Activation

Roediger, Balota, and Watson (2001) suggest that the DRM effect is due to automatic associative activation and mainly focus on word-based similarity or associations, but in the current study, word-based similarity cannot fully account for the results. In fact, the critical and control lures did not differ significantly in terms of word similarity to their respective study lists, at least as measured in terms of a vector-based word model, which should tap into both similarity and association (Clark, 2015; Reid & Katz, 2018b). Therefore, it seems that false memory in the DRM paradigm can be elicited by factors beyond word similarity. In this case, it was elicited by metaphorical similarity.

Automatic spreading activation could accommodate the current findings if metaphorical expressions are stored in semantic memory somewhat like words (e.g., Swinney & Cutler, 1979; see Nayak & Gibbs, 1990, for a review of these accounts). Similar to how nodes for related words are connected, nodes for expressions might also be connected when they are based on the same underlying conceptual metaphor. That is, expressions that use the same cross-domain mapping are likely stored in the same semantic memory structure, or at least these expressions are in closer proximity than expressions using different cross-domain mappings. By this account, activation could possibly spread between nodes for expressions as it does with nodes for words. I repeat this is speculation because expressions really have not been considered in the spreading activation literature, especially regarding DRM false memory effects.

### 6.1.2 Processing Fluency

Processing fluency and spreading activation explanations of false memory go hand-in-hand as spreading activation results in fluent processing of related lures (Gallo & Roediger, 2003; Whittlesea, 2002). However, the critical difference is that processing fluency does not depend on metaphorical expressions being stored in semantic memory. Alternatively, conceptual metaphors may act more like schemas that help to organize and interpret metaphorical expressions and the schema may be activated when a metaphorical expression is read. In other words, the expression itself is not retrieved from semantic memory, but the conceptual metaphor schema is employed to make sense of the expression. This would still lead to the same processing fluency advantage because if the schema is already activated from reading prior expressions, new expressions should be more readily processed.

A processing fluency explanation could be tested using the paradigm employed by Whittlesea (2002), who had participants read DRM-type lists, but then make a lexical decision on the critical lure rather than a recognition judgment. Applied to the current study, participants would first read the study list expressions and then read related and unrelated critical consistent lure sentences. A processing fluency explanation would depend on the critical consistent lures being read faster when related to the study list than when unrelated. Some previous research has found that reading metaphor expressions based on the same conceptual metaphor facilitates processing of related expressions, but the findings are mixed, and either way, facilitated processing has not been confirmed with the stimuli employed in the current study.

### 6.1.3 Fuzzy-trace theory

Lastly, the conceptual metaphor false memory effect could be explained in terms of fuzzy-trace theory (Brainerd & Reyna, 2005). Recall that this theory focuses on the "gist," or the deeper theme or meaning of the to-be-remembered information. From this perspective, the "gist" of each study list would essentially be the underlying conceptual metaphor, and the critical consistent lures would be consistent with the "gist" of each list. Although some have argued that gist extraction likely relies on more effortful processing, and therefore, more conscious awareness than a spreading activation explanation (Carneiro, Garcia-Marques, Fernandez, & Albuquerque, 2014), Brainerd and colleagues (Brainerd & Reyna, 2005; Brainerd, Forrest, Karibian, & Reyna, 2006) do not make any claims about whether gist extraction is conscious or unconscious. Therefore, a gist-based explanation for the findings from this set of studies accommodates the data nicely.

The purpose of this thesis was not to test alternative theories of false memory, but to test assumptions of CMT, so I do not make strong claims regarding which memory theory best supports the data. Overall, the three outlined theories all reasonably accommodate the findings.

### 6.1.4 Remember, Know, and Guess Judgments

Across Chapters 3, 4, and 5 in which the remember-know-guess procedure was employed, the results suggest that false recognitions of the critical consistent lures were attributed mostly to familiarity. Over all these chapters, the falsely recognized critical consistent lures never had a percentage of remember judgments above 50%, and the percentage of guess judgments was fairly high across all experiments, suggesting a weaker and less confident sense of familiarity (Hirshman, 1998). Although these data on their own cannot be used to confidently distinguish between CMT and DMT, when considered in light of the other findings from the strategy descriptions and divided attention manipulations, they further suggest that conceptual metaphors are activated automatically. Familiarity is considered to be the more unconscious and automatic type of memory whereas recollection is considered more effortful and intentional. Given the automaticity assumption of CMT, one would expect a greater degree of familiarity-based false recognitions. This is also consistent with the notion that conceptual metaphor activation facilitates processing of related expressions, as processing fluency primarily affects familiarity. In contrast, one may expect more false recollection if cross-domain mappings depend on conscious and deliberate processing, as proposed by DMT. I repeat that this cannot carry the argument for CMT over DMT; however, given the entirety of the data, the high percentage of false familiarity further supports the automaticity of conceptual metaphor activation.

It should be noted that the low percentage of “remember” responses could also be due to methodological factors unrelated to CMT and DMT. For instance, full sentences were employed, and participants may have hesitated to use “remember” judgments if any single word in the lure sentence seemed unfamiliar. Full sentences also require a longer presentation duration (3 seconds) which is a factor known to weaken false recognition effects (McDermott & Watson, 2001) and potentially resulted in less false “remember” judgments in the current study. At best, the “remember,” “know,” and “guess” judgments should be interpreted cautiously, but the data is consistent with the original CMT.

## 6.2 Implications for CMT research

These experiments provide a novel way of testing CMT and have demonstrated that conceptual metaphors play a role in episodic memory. As Murphy (1996) pointed out in his critique rejecting the claim that CMT is a realistic theory of conceptual representation, the influence of conceptual metaphors should be observable across the broad domains of cognitive psychology, such as memory, problem solving, and categorization. He rightfully pointed out that CMT research has not systematically explored these different areas. The studies reported here have extended the study of CMT into one of the domains proposed as important by Murphy, namely the implications of CMT on episodic memory tasks.

To date, most of the cognitive (psychological) research on CMT falls under one of three broad categories: embodiment of concepts, metaphorical framing, and online comprehension. Each of these streams of research is valuable and has demonstrated compelling effects of metaphor in thought, though as reviewed below, each has issues for being a pure measure of the influence of conceptual metaphors on cognition.

In terms of embodiment, it is not unique to CMT. Barsalou and Wiemer-Hastings (2005) argue that abstract concepts can be embodied through direct experience rather than metaphorical mappings, and Holyoak and Stamenković (2018) argue embodied effects could be explained by polysemous words activating both physical embodied and abstract senses of words simultaneously. Although embodiment has been a fundamental aspect of CMT, some scholars have challenged the more rigid view that conceptual metaphors are always concrete source-to-abstract target mappings. Ortony (1988) argued that for many emotion metaphors (e.g., ANGER IS HEAT IN A CONTAINER, ANGER IS INSANITY), children would have experience with the emotion before they would have an understanding of the source domain that structures it. Source domains are presumed to supply information from easier understood and directly experienced concepts, yet to understand the various ANGER IS HEAT IN A CONTAINER metaphors (e.g., “*he* *flipped his lid*"), one would need to know at least some information about physics (see also Murphy, 1996; Thibodeau & Durgin, 2008). However, it is unlikely that children's knowledge of physics would precede their knowledge of anger. Therefore, in this case it seems that the target domain is the more easily understood and directly experienced domain, especially early in life. Gibbs (2018) has also argued that CMT has neglected abstract source-to-concrete target mappings, for instance, not only can JOURNEYS be used to conceptualize LIFE, but LIFE can also be a metaphor for a JOURNEY. Therefore, although many conceptual metaphors may help to conceptualize abstract concepts by drawing on concrete experiences, it seems that this need not always be the case; conceptual metaphors may simply draw information from another domain that in some way helps to highlight or communicate certain aspects of the target domain (Thibodeau & Durgin, 2008). As such, CMT research exploring the embodiment of abstract concepts may not capture the full spectrum of how conceptual metaphors are used in thought.

The stream of research on metaphorical framing has emphasized ecologically valid experiments exploring how metaphors can influence reasoning about real-life issues, such as crime (Steen, Reijnierse, & Burgers, 2014; Thibodeau & Boroditsky, 2011, 2013), marriage (Robins & Mayer, 2000), and global warming (Flusberg, Matlock, & Thibodeau, 2017). It should be noted that this line of research is not primarily concerned with testing the assumptions of CMT, but rather the focus is on the metaphorical framing effect itself. However, metaphorical framing aligns nicely with CMT as the fundamental assumption of CMT is that metaphors are a matter of thought, not just of language.

As a pure measure of the effect of conceptual metaphors on cognition, metaphorical framing has some issues. For instance, the metaphorical framing effect disappears when the participant is already an expert on the target domain being framed (Robins & Mayer, 2000). Also, with social and political issues, participants tend to bring in many of their own views and beliefs that are not necessarily malleable, and certain individual differences influence the strength of the metaphorical framing effect. For instance, Thibodeau and Boroditsky (2011) found that, with respect to politics of the USA, self-identified Republicans were more resistant to metaphorical framing than Democrats and

Independents. Arguably, the episodic memory task used here is less influenced by such social factors. Lastly, with metaphorical framing, it is not always clear what type of reasoning is congruent with what type of metaphor. For instance, two different metaphors could lead to similar inferences. With CRIME IS A VIRUS/BEAST frames, “locking up” criminals is argued to be congruent with the BEAST frame as it emphasizes punishment (Thibodeau & Boroditsky, 2011). However, viral outbreaks often involve quarantining infected individuals to protect the rest of the public, such as has been seen with the coronavirus currently. Therefore, “locking up” criminals could be considered an inference following from the CRIME IS A VIRUS frame. Again, such interpretive factors are not an issue in the episodic memory tests employed here. This is not to say that metaphorical framing research lacks value; however, for the purpose of testing CMT itself, episodic memory tests have certain advantages.

The last stream of research that most directly assesses CMT is online comprehension, which has generated mixed findings. Though online comprehension is a logical variable to explore, it may not be the only area in which conceptual metaphors influence cognition. In fact, Bundgaard (2019) has recently proposed that cross-domain mappings are psychologically real, but that their activation is not required to comprehend conventional metaphor expressions. Conceptual metaphors may not be strictly required for comprehension, but they may serve as useful schemas that help to interpret, organize and encode metaphorical language, especially when the same cross-domain mapping is extended over several different expressions. If this is the case, this may be why effects supportive of CMT are found more consistently in offline studies (Nayak & Gibbs, 1990; Katz & Law, 2010; Katz & Taylor, 2008, and studies reported here) than in online studies.

One other possible advantage of the current study over the previous comprehension studies is that it employed sentences that were putatively unrelated. In contrast, the comprehension studies mentioned earlier used narratives in which the sentences all logically connect. The issue here is that the narrative itself may influence online comprehension. For instance, Thibodeau and Durgin (2008) point out that comprehension times may be affected by the relation of the critical sentence to the meaning of the preceding narrative, and not only by the conceptual metaphor. This is not an issue in the current study as the sentences were unrelated on the surface.

## 6.3 Implications for Deliberate Metaphor Theory

The findings from this set of experiments do not support Deliberate Metaphor Theory. According to this theory, cross-domain mappings that would lead to the false memory effects found in these experiments are proposed to depend on conscious awareness of metaphoricity. When a metaphor is deliberate, that is, the speaker is intentionally using the metaphor as a metaphor, the deliberateness is thought to afford conscious metaphorical cognition. In contrast, non-deliberate metaphors do not afford conscious metaphorical cognition, and therefore, do not activate cross-domain mappings.

It should be noted that these experiments did not manipulate deliberateness, unlike Gibbs's (2015) test of DMT in which deliberateness was manipulated via "pragmatic markers" (words such as "like" that putatively signal a metaphorical comparison). Gibbs found no evidence that the pragmatic markers influenced metaphor interpretation. However, he only examined a single metaphorical utterance ("We really have come a long way since the wedding") and asked participants about the meaning of the utterance, which may have in itself drawn deliberate attention to its metaphoricity. I am not arguing that the concept of metaphor deliberateness be discounted as playing any role on metaphor studies, especially in communicative contexts. It is possible that deliberateness enhances conceptual metaphor activation, above activation occurring unconsciously, and could possibly elicit stronger effects on memory.[17] That possibility aside, Deliberate

[17] In fact, one could test the effect of deliberateness on conceptual metaphor activation using the current paradigm. For instance, study lists consisting of deliberate metaphor expressions could be compared to study lists consisting of non-deliberate metaphor expressions. If the deliberate study lists induce a greater proportion of false recognition for the critical consistent lures, it would suggest that deliberateness enhances conceptual metaphor activation. Similar between list comparisons have been conducted in other DRM research, such as lists based on associates vs. categories (Buchanan, Brown, Cabeza, & Maitson, 1999; Knott & Dewhurst, 2007), situation lists vs. DRM lists (Cann, McRae, & Katz, 2011, Experiment 2), and lists varying on gist-strength (Cann, McRae, & Katz, 2011, Experiment 4).

Metaphor Theory explicitly states that deliberateness affords conscious metaphorical cognition, and it is consciousness that is the critical factor for a cross-domain mapping to be engaged. The current data strongly refutes this claim.

## 6.4 Conclusion

This set of studies employed a novel technique to test the tenets of CMT and found strong evidence that conceptual metaphors are psychologically real and influence cognition automatically. By using the DRM paradigm, I have tested assumptions of CMT in terms of a well-established, robust memory task familiar to cognitive psychologists. Across all the studies, it was found that studying a list of expressions all based on the same conceptual metaphor led to false recognitions of other expressions that were also based on the same conceptual metaphor. These expressions were falsely recognized more often than control expressions that also shared semantic overlap with the study list, but were not based on the same metaphor. Also, converging evidence from self-reported strategies (Chapter 3) and divided attention manipulations (Chapters 4 and 5) indicated that the conceptual metaphors were engaged automatically, as participants rarely reported consciously attending to the conceptual metaphors and still showed the effect even when their attention was divided at either encoding or retrieval (or both). As such, this set of studies supports the original conception of CMT proposed by Lakoff and Johnson (1980) and Lakoff (1993). These studies also emphasize the utility of episodic memory tasks for exploring tenets of CMT, and future research could employ other episodic tasks to explore the boundary conditions under which conceptual metaphors organize and influence memory.

# References

Allbritton, D. W., McKoon, G., & Gerrig, R. J. (1995). Metaphor-based schemas and text representations: Making connections through conceptual metaphors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(3), 612-625.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22,* 577-660.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617-645.

Barsalou, L. W., & Wiemer-Hastings, K. (2005). Situating abstract concepts. In D. Pecher and R. A. Zwaan (Eds.), *Grounding cognition: The role of perception and action in memory, language and thinking* (pp. 129-163). Cambridge, England: Cambridge University Press.

Bousfield, W. A. (1953). The occurrence of clustering in the recall of randomly arranged associates. *The Journal of General Psychology*, *49*(2), 229-240.

Bowdle, B. F., & Gentner, D. (2005). The career of metaphor. *Psychological Review, 112*(1), 193-216.

Brainerd, C. J., Forrest, T. J., Karibian, D., & Reyna, V. F. (2006). Development of the false-memory illusion. *Developmental Psychology*, *42*(5), 962–979.

Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science*, *11*(5), 164-169.

Brainerd, C. J., & Reyna, V. F. (2005). *The science of false memory.* New York, NY: Oxford University Press.

Brainerd, C. J., Reyna, V. F., & Mojardin, A. H. (1999). Conjoint recognition. *Psychological Review*, *106*(1), 160–179.

Bransford, J. D., & Franks, J. J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology*, *2*(4), 331-350.

Buchanan, L., Brown, N. R., Cabeza, R., & Maitson, C. (1999). False memories and semantic lexicon arrangement. *Brain and Language*, *68*(1-2), 172-177.

Bundgaard, P. F. (2019). The structure of our concepts: A critical assessment of Conceptual Metaphor Theory as a theory of concepts. *Cognitive Semiotics*, *12*(1).

Cann, D. R., & Katz, A. N. (2005). Habitual acceptance of misinformation: Examination of individual differences and source attributions. *Memory & Cognition*, *33*(3), 405-417.

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*(8), 1515-1542.

Cardillo, E. R., Schmidt, G. W., Kranjec, A., & Chatterjee, A. (2010). Stimulus design is an obstacle course: 560 matched literal and metaphorical sentences for testing neural hypotheses about metaphor. *Behavior Research Methods*, *42*(3), 651-664.

Carneiro, P., Garcia-Marques, L., Fernandez, A., & Albuquerque, P. (2014). Both associative activation and thematic extraction count, but thematic false memories are more easily rejected. *Memory, 22*(8), 1024-1040.

Casasanto, D. (2008). Similarity and proximity: When does close in space mean close in mind? *Memory & Cognition, 36*(6), 1047–1056

Clark, S. (2015). Vector space models of lexical meaning. In S. Lappin & C. Fox (Eds.), *The Handbook of Contemporary Semantic Theory, 2nd ed.* (pp. 493-522). Malden, MA: Wiley-Blackwell.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*(1), 17-22.

Desai, R. H., Binder, J. R., Conant, L. L., Mano, Q. R., & Seidenberg, M. S. (2011). The neural career of sensory-motor metaphors. *Journal of Cognitive Neuroscience, 23*(9), 2376-2386.

Dewhurst, S. A., Barry, C., Swannell, E. R., Holmes, S. J., Bathurst, G. L. (2007). The effect of divided attention on false memory depends on how memory is tested. *Memory & Cognition, 35*(4), 660-667.

Doss, M. K., Bluestone, M. R., & Gallo, D. A. (2016). Two mechanisms of constructive recollection: Perceptual recombination and conceptual fluency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(11), 1747–1758.

Faul, F., Erdfelder, E., Lang, A., Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191.

Flusberg, S. J., Matlock, T., & Thibodeau, P. H. (2017). Metaphors for the war (or race) against climate change. *Environmental Communication*, *11*(6), 769-783.

Fusaroli, R. & Morgagni, S. (Eds.). (2009). Conceptual metaphor theory: Thirty years after. *The Journal of Cognitive Semiotics*, *5*(1-2).

Gallo, D. A. (2006). *Associative illusions of memory: False memory research in DRM and related tasks.* New York, NY: Psychology Press.

Gallo, D. A., Roberts, M. J., & Seamon, J. G. (1997). Remembering words not presented in lists: Can we avoid creating false memories? *Psychonomic Bulletin & Review, 4*(2), 271-276.

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*(3), 469-497.

Gallo, D. A., & Roediger, H. L. (2003). The effects of associations and aging on illusory recollection. *Memory & Cognition, 31*(7), 1036-1044.

Gallo, D. A., Roediger, H. L., & McDermott, K. B. (2001). Associative false recognition occurs without strategic criterion shifts. *Psychonomic Bulletin & Review, 8*(3), 579-586.

Gardiner, J. M., (1988). Functional aspects of recollective experience. *Memory & Cognition, 16*(4), 309-313.

Gardiner, J. M., Ramponi, C., & Richardson-Klavehn, A. (1998). Experiences of remembering, knowing, and guessing. *Consciousness and Cognition, 7*(1), 1-26.

Gardiner, J. M., Craik, F. I. M., & Birtwistle, J. (1972). Retrieval cues and release from proactive inhibition. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 778-783.

Gibbs, R. W. (1996). Why many concepts are metaphorical. *Cognition, 61*(3), 309-319.

Gibbs, R. W. (2006). *Embodiment and cognitive science.* New York, NY: Cambridge University Press.

Gibbs, R. W. (2009). Why do some people dislike conceptual metaphor theory? *The Journal of Cognitive Semiotics, 5*(1-2), 14-36.

Gibbs, R. W. (2011). Evaluating Conceptual Metaphor Theory. *Discourse Processes, 48*(8), 529-562.

Gibbs, R. W. (2013). The real complexities of psycholinguistic research on metaphor. *Language Sciences, 40*, 45-52.

Gibbs, R. W. (2015). Do pragmatic signals affect conventional metaphor understanding? A failed test of deliberate metaphor theory. *Journal of Pragmatics, 90,* 77-87.

Gibbs, R. W. (2018, June). *A little taste of heaven: Metaphor in bodily experience.* Talk presented at the 12th Researching and Applying Metaphor Conference, Hong Kong. Abstract available from http://www.engl.polyu.edu.hk/events/raam/book%20of%20abstracts.pdf

Glucksberg, S., Brown, M., & McGlone, M. S. (1993). Conceptual metaphors are not automatically accessed during idiom comprehension. *Memory & Cognition, 21*(5), 711-719.

Glucksberg, S., & Keysar, B. (1990). Understanding metaphorical comparisons: Beyond similarity. *Psychological Review, 97*(1), 3-18.

Glucksberg, S., & McGlone, M. S. (1999). When love is not a journey: What metaphors mean. *Journal of Pragmatics, 31*(12), 1541-1558.

Gong, S.-P., & Ahrens, K. (2007). Processing conceptual metaphors in on-going discourse. *Metaphor and Symbol, 22*(4), 313-330.

Hirshman, E. (1998). On the utility of the signal detection model of the remember–know paradigm. *Consciousness and Cognition: An International Journal, 7*(1), 103-107.

Holyoak, K. J., & Stamenković, D. (2018). Metaphor comprehension: A critical review of theories and evidence. *Psychological Bulletin, 144*(6), 641–671.

Hyönä, J. (1993). Effects of thematic and lexical priming on readers' eye movements. *Scandinavian Journal of Psychology, 34*(4), 293-304.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional used of memory. *Journal of memory and language, 30*(5), 513-541.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 110*(3), 306–340.

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

Jacoby, L. L., & Whitehouse, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology: General, 118*(2), 126–135.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114*(1), 3-28.

Katz, A., & Law, A. (2010). Experimental support for conceptual metaphors with an episodic memory task. *Metaphor and Symbol*, *25*(4), 263-270.

Katz, A. & Reid, J. N. (2020). Tests of Conceptual Metaphor Theory with episodic memory tests. *Cognitive Semantics, 6*(1), 56-82.

Katz, A. N., & Taylor, T. E. (2008). The journeys of life: Examining a conceptual metaphor with semantic and episodic memory recall. *Metaphor and Symbol, 23*(3), 148-173.

Kawasaki, Y. & Yama, H. (2006). The difference between implicit and explicit associative processes at study in creating false memory in the DRM paradigm. *Memory, 14*(1), 68-78.

Kelley, C. M., & Jacoby, L. L. (1998). Subjective reports and process dissociation: Fluency, knowing, and feeling. *Acta Psychologica*, *98*(2-3), 127-140.

Keysar, B., Shen, Y., Glucksberg, S., & Horton, W. S. (2000). Conventional language: How metaphorical is it? *Journal of Memory and Language*, *43*(4), 576-593.

Knott, L. M., & Dewhurst, S. A. (2007). The effects of divided attention at study and test on false recognition: A comparison of DRM and categorized lists. *Memory & Cognition*, *35*(8), 1954-1965.

Lakoff, G. (1993). The contemporary theory of metaphor. In A. Ortony (Ed.), *Metaphor and thought: Vol. 2* (pp. 202–251). New York, NY: Cambridge University Press.

Lakoff, G. (1994). *The Master Metaphor List.* Retrieved from http://www.lang.osaka-u.ac.jp/~sugimoto/MasterMetaphorList/MetaphorHome.html

Lakoff, G. (2008). The neural theory of metaphor. In R. W. Gibbs (Ed.), *Cambridge handbook of metaphor and thought* (pp. 17–38). New York, NY: Cambridge University Press.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes, 25*(2-3), 259-284.

Leboe, J. P., & Whittlesea, B. W. A. (2002). The inferential basis of familiarity and recall: Evidence for a common underlying process. *Journal of Memory and Language, 46*(4), 804-829.

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory, 4*(1), 19-31.

McDermott, K. B., & Roediger, H. L. (1998). Attempting to avoid illusory memories: Robust false recognition of associates persists under conditions of explicit warnings and immediate testing. *Journal of Memory and Language, 39*(3), 508-520.

McDermott, K. B., & Watson, J. M. (2001). The rise and fall of false recall: The impact of presentation duration. *Journal of Memory and Language, 45*(1), 160-176.

McGlone, M. S. (2007). What is the explanatory value of a conceptual metaphor? *Language & Communication, 27*(2), 109-126.

McGlone, M. S. (2011). Hyperbole, homunculi, and hindsight bias: An alternative evaluation of conceptual metaphor theory. *Discourse Processes, 48*(8), 563-574.

Multhaup, K. S., & Conner, C. A. (2002). The effects of considering nonlist sources on the Deese–Roediger–McDermott memory illusion. *Journal of Memory and Language, 47*(2), 214-228.

Murphy, G. L. (1996). On metaphoric representation. *Cognition, 60*(2), 173-204.

Nayak, N. P., & Gibbs, R. W. (1990). Conceptual knowledge in the interpretation of idioms. *Journal of Experimental Psychology: General, 119*(3), 315–330.

Ortony, A. (1988). Are emotion metaphors conceptual or lexical? [Review of the book *Metaphors of anger, pride and love: A lexical approach to the structure of concepts,* by Z. Kövecses]. *Cognition and Emotion, 2*(2), 95-103.

Otgaar, H., Peters, M., & Howe, M. L. (2012). Dividing attention lowers children's but increases adults' false memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(1), 204–210.

Park, L., Shobe, K. K., & Kihlstrom, J. F. (2005). Associative and categorical relations in the associative memory illusion. *Psychological Science, 16*(10), 792-797.

Pérez-Mata, M. N., Read, J. D., & Diges, M. (2002). Effects of divided attention and word concreteness on correct recall and false memory reports. *Memory, 10*(3), 161-177.

Pimentel, E., & Albuquerque, P. B. (2013). Effect of divided attention on the production of false memories in the DRM paradigm: A study of dichotic listening and shadowing. *Psicológica, 34*(2), 285-298.

Posner, M. I., & Keele, S. W. (1970). Retention of abstract ideas. *Journal of Experimental Psychology, 83*(2), 304-308.

Rajaram, S. (1993). Remembering and knowing: Two means of access to the personal past. *Memory & Cognition, 21*(1), 89-102.

Rajaram, S., & Geraci, L. (2000). Conceptual fluency selectively influences knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(4), 1070–1074.

Reid, J. N., & Katz, A. N. (2018a). Something false about conceptual metaphors. *Metaphor and Symbol, 33*(1), 36-47.

Reid, J. N., & Katz, A. N. (2018b). Vector space applications in metaphor comprehension. *Metaphor and Symbol, 33*(4), 280-294.

Reijnierse, W. G., Burgers, C., Krennmayr, T., & Steen, G. J. (2018). DMIP: A method for identifying potentially deliberate metaphor in language use. *Corpus Pragmatics, 2*, 129-147.

Renoult, L., Irish, M., Moscovitch, M., & Rugg, M. D. (2019). From knowing to remembering: The semantic-episodic distinction. *Trends in Cognitive Sciences, 23*(12), 1041-1057.

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences, 7*(1), 1-75.

Robins, S., & Mayer, R. E. (2000). The metaphor framing effect: Metaphorical reasoning about text-based dilemmas. *Discourse Processes, 30*(1), 57-86.

Roediger, H. L., Balota, D. A., & Watson, J. M. (2001). Spreading activation and arousal of false memories. In H. L. Roediger, J. S. Nairne, I. Neath, & A. M. Surprenant (Eds.), *Science conference series. The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 95–115). 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*(4), 803-814.

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*(3), 385-407.

Schacter, D. L., Verfaellie, M., & Pradere, D. (1996). The neuropsychology of memory illusions: False recall and recognition in amnesic patients. *Journal of Memory and Language, 35*(2), 319-334.

Seamon, J. G., Luo, C. R., & Gallo, D. A. (1998). Creating false memories of words with or without recognition of list items: Evidence for nonconscious processes. *Psychological Science, 9*(1), 20-26.

Steen, G. (2008). The paradox of metaphor: Why we need a three-dimensional model of metaphor. *Metaphor and Symbol, 23*(4), 213-241.

Steen, G. (2009). Deliberate metaphor affords conscious metaphorical cognition. *Cognitive Semiotics, 5*(1-2), 179-197.

Steen, G. (2011). What does 'really deliberate' really mean?: More thoughts on metaphor and consciousness. *Metaphor and the Social World, 1*(1), 53-56.

Steen, G. (2015). Developing, testing and interpreting deliberate metaphor theory. *Journal of Pragmatics, 90,* 67-72.

Steen, G. (2017). Deliberate Metaphor Theory: Basic assumptions, main tenets, urgent issues. *Intercultural Pragmatics*, *14*(1), 1-24.

Steen, G. J., Reijnierse, W. G., Burgers, C. (2014). When do natural language metaphors influence reasoning? A follow-up study to Thibodeau and Boroditsky (2013). *PLoS One*, *9*(12), e113536.

Swinney, D. A., & Cutler, A. (1979). The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behavior*, *18*(5), 523-534.

Thibodeau, P. H., & Boroditsky, L. (2011). Metaphors we think with: The role of metaphor in reasoning. *PLoS ONE, 6*(2), e16782.

Thibodeau, P. H., & Boroditsky, L. (2013). Natural language metaphors covertly influence reasoning. *PLoS ONE, 8*(1), e52961.

Thibodeau, P. H., & Durgin, F. H., (2008). Productive figurative communication: Conventional metaphors facilitate the comprehension of related novel metaphors. Journal of Memory and Language, 58(2), 521-540.

Traxler, M. J., Foss, D. J., Seely, R. E., Kaup, B., & Morris, R. K. (2000). Priming in sentence processing: Intralexical spreading activation, schemas, and situation models. *Journal of Psycholinguistics Research, 29,* 581-595.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of Memory* (pp. 381-403). Oxford, England: Academic Press.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology/Psychologie canadienne*, *26*(1), 1-12.

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

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*(5), 352-373.

Watkins, O. C., & Watkins, M. J. (1975). Buildup of proactive inhibition as a cue-overload effect. *Journal of Experimental Psychology: Human Learning and Memory*, *1*(4), 442–452.

Weidemann, C. T., Kragel, J. E., Lega, B. C., Worrell, G. A., Sperling, M. R., Sharan, A. D., Jobst, B. C., Khadjevand, F., Davis, K. A., Wanda, P. A., Kadel, A., Rizzuto, D. S., & Kahana, M. J. (2019). Neural activity reveals interactions between episodic and semantic memory systems during retrieval. *Journal of Experimental Psychology: General*, *148*(1), 1–12.

Whittlesea, B. W. A. (2002). False memory and the discrepancy-attribution hypothesis: The prototype-familiarity illusion. *Journal of Experimental Psychology: General, 131*(1), 96–115.

Wickens, D. D. (1970). Encoding categories of words: An empirical approach to meaning. *Psychological Review*, *77*(1), 1-15.

Yonelinas, A. P. (2001). Consciousness, control, and confidence: The 3 Cs of recognition memory. *Journal of Experimental Psychology: General, 130*(3), 361-379.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46*(3), 441-517.

# Appendices

## Appendix A: Copyright permission email

February 12, 2020

Dear Professor Albert Katz on Behalf of J. Nick Reid,

**Material Requested: J. Nick Reid & Albert N. Katz (2018)**

**Something false about conceptual metaphors,**

***Metaphor and Symbol*, 33:1, 36-47, DOI: 10.1080/10926488.2018.1407994**

Thank you for your correspondence requesting permission to post the 'Accepted Manuscript' of your above mentioned article on **your institutions intranet or within the Institutional Repository of your institution and in your dissertation.**

As you are the author of the above article, we will be pleased to grant entirely free permission on the condition that you make the following acknowledgement:

"This is an **Accepted Manuscript** of an article published by Taylor & Francis in ***Metaphor and Symbol*** on **January 3, 2018**, available online at the Taylor & Francis Ltd web site: www.tandfonline.com and include a link to the article - **https://www.tandfonline.com/doi/full/10.1080/10926488.2018.1407994**

You will also need to obtain permission from any co-authors of this article.

Please note that this license does not allow you to post our content on any other third party websites.

This license does not allow the use of the Publishers version/PDF (this is the version of record that is published on the publisher's website) to be posted online.

Thank you for your interest in our Journals.

Sincerely,

**Appendix B: Study lists and lures**

IDEAS ARE FOOD study list

What he said left a bad taste in my mouth.
That plan is on the back burner for now.
All this paper has are raw facts.
Learning is eating.
Those are warmed-over arguments.
We have to regurgitate everything we learned on the final.
Here's a concept you can sink your teeth into.
The plan is half-baked.
This is the meaty part of the paper.
He devoured the book.
Her curiosity is insatiable.
I'll give you some readings to chew on.
Knowledge is consumable.
She cooked up a new scheme.
We don't spoon-feed our students.

Recognition test lures

(Critical CM)
-Ideas are food.

(Critical consistent)
-That claim is hard to swallow.
-He has an appetite for learning.
-There are too many facts to digest. (Final test only)

(Control metaphor)
-That kind of thinking is out of style.
-His first lecture just planted the seeds.
-She never arrives at the right conclusion. (Final test only)

(Control literal)
-We discussed the plan over dinner.
-The dessert was too sweet.
-This dish is best served cold. (Final test only)
-I came up with this scheme during lunch. (Final test only)

<u>LOVE IS A JOURNEY study list</u>

They're at a crossroads in their relationship.
This marriage is on the rocks.
We've come a long way as a couple.
Dating is a starting point.
It seems we're just going in circles as a couple.
They're in a dead-end relationship.
We've had some bumps in the road.
Relationships are vehicles.
My girlfriend and I may have to go our separate ways.
Where are we in this relationship?
Marriage is a landmark.
My fiancé and I can't turn back now.
Break-ups are obstacles.
Their marriage has gone off the track.
I want to take things slow in this relationship.

<u>Recognition test lures</u>

(Critical CM)
-Love is a journey.

(Critical consistent)
-Their romance just took a turn for the worst.
-They didn't take the path most couples take.
-My boyfriend and I are stuck in a rut. (Final test only)

(Control metaphor)
-Their relationship has lost its magic.
-She's crazy about him.
-She swept him off his feet. (Final test only)

(Control literal)
-I took my girlfriend on a nice drive.
-We took a short-cut to our destination.
-You can only get to this place on foot. (Final test only)
-This pathway is a short walk. (Final test only)

<u>THEORIES ARE BUILDINGS study list</u>

Your argument had a nice structure.
Our method will stand or fall on the strength of that claim.
His assumptions are built on sand.
Scientists are architects.
What will the form of the argument be?
They demolished his reasoning.
Without data our model will fall apart.
Evidence is support.
The argument collapsed.
That’s a shaky assumption.
Research is construction.
We need to buttress our paper with solid facts.
Facts are foundation.
She tore down his argument brick by brick.
Here are some more data to prop up the hypothesis.

<u>Recognition test lures</u>

(Critical CM)
-Theories are buildings.

(Critical consistent)
-That claim doesn’t hold much weight.
-Her work was a pillar in the discipline.
-We have put together only the framework of this hypothesis. (Final test only)

(Control metaphor)
-His argument fell apart at the seams.
-That paper gave birth to new lines of research.
-Her proposition never bore any fruit. (Final test only)

(Control literal)
-His house was well constructed.
-His apartment had an interesting layout.
-The centre has multiple levels. (Final test only)
-Her condo has three rooms. (Final test only)

TIME IS MONEY study list

How did you spend the summer break?
I have some days off banked from last month.
Budget your hours.
Weekends are precious.
I don't have the hours for this.
I'll give you a minute.
Is that worth your while?
Years are invested.
Put aside a few days for this.
Can you spare an afternoon?
Hours are wasted.
How many minutes do I have left?
Free hours are valuable.
The diversion should buy him a few minutes.
This will save me many hours.

TIME IS MONEY recognition test lures

(Critical CM)
-Time is money

(Critical consistent)
-Lend me a few minutes.
-That cost me a day.
-You don't use your hours profitably. (Final test only)

(Control metaphor)
-The weekend seems so far away.
-The years have not been kind to him.
-The deadline is approaching. (Final test only)

(Control literal)
-How much is your rent per month?
-He makes biweekly payments.
-She took out a low-interest loan. (Final test only)
-I will pay you back in a week. (Final test only)

UNDERSTANDING IS SEEING study list

Here's another way you can look at this problem.
Before you respond, let me first point something out.
That was an insightful dialogue.
The truth is clear.
Her thoughts on the subject are muddy.
With this issue, you have to look at the whole picture.
That was a brilliant remark.
Explaining is illuminating.
It was a murky discussion.
Could you elucidate your remarks?
Falseness is darkness.
It's a transparent argument.
Ignorance is blindness.
The discussion was opaque.
He has tunnel-vision when it comes to this issue.

UNDERSTANDING IS SEEING recognition test lures

(Critical CM)
-Understanding is seeing.

(Critical consistent)
-We have different views on this subject.
-That lecture opened my eyes.
-Let me enlighten you on this topic. (Final test only)

(Control metaphor)
-I couldn't grasp his argument.
-That lesson was in one ear and out the other.
-It took a while, but the concept finally clicked. (Final test only)

(Control literal)
-Her vision is blurry.
-The lack of lighting caused low visibility.
-I have a nice view from my office. (Final test only)
-I think I can spot my house from here. (Final test only)

**Appendix C: Remember, know, and guess instructions. Adjusted from Gardiner, Ramponi, and Richardson-Klavehn (1998).**

PLEASE READ THE FOLLOWING INSTRUCTIONS CAREFULLY.

Recognition memory is associated with two different kinds of awareness. Quite often recognition brings back to mind something you recollect about what it is that you recognise, as when, for example, you recognize someone’s face, and perhaps remember talking to this person at a party the previous night. At other times recognition brings nothing back to mind about what it is you recognise, as when, for example, you are confident that you recognise someone, and you know you recognise them, because of strong feelings of familiarity, but you have no recollection of seeing this person before. You do not remember anything about them.

The same kinds of awareness are associated with recognising the sentences on the study lists. Sometimes when you recognize a sentence as one you saw on the study list, recognition will bring back to mind something you remember thinking about when the sentence appeared then. You recollect something you consciously experienced at that time. But sometimes recognizing a sentence as one you saw on the study list will not bring back to mind anything you remember about seeing it then. Instead, the sentence will seem familiar, so that you feel confident it was one you saw from the study list, even though you don’t recollect anything you experienced when you saw it then.

For each sentence that you recognize, after you have pressed “O” (for OLD), please then press “R” (for REMEMBER), if recognition is accompanied by some recollective experience, or “K” (for KNOW), if recognition is accompanied by strong feelings of familiarity in the absence of any recollective experience.

There will also be times when you do not remember the sentence, nor does it seem familiar, but you might want to guess that it was one of the sentences you saw on the study list. Feel free to do this, but if your OLD response is really just a guess, please press “G” (for GUESS).

If you have any questions regarding these judgments, please ask the experimenter. If you need to be reminded of what these judgments mean during the experiment, please refer to the sheet on the desk that has these instructions printed out. Thank you.

Please press the REMEMBER key to continue.

**Appendix D: Strategy instructions for the study presented in Chapter 3.**

Could you please take a minute to describe any strategies you used to remember this list? If you used a strategy, please provide as much information about your strategy as you can. Really, there are no right or wrong answers; we are just trying to get insight into how people try to remember sentences. If you did not use any particular strategy, just type in: "I just tried to remember the sentences without using any conscious strategy."

Here are some possible strategies you might have used:

- repeating the sentences over and over mentally (if so, please estimate whether you repeated the sentences very frequently, frequently, very infrequently),

- looking for a theme in the sentences presented that you thought might help you remember the sentences (if so, please indicate the theme),

- trying to connect the sentences to a mental image or sound pattern (if so, please tell us as much about the image or sound pattern as possible),

- trying to focus on specific words that you thought were very memorable (if so, please indicate these words),

- some other strategy. If so, please briefly describe what it might be.

Please type your response below.

When you are finished typing your answer, press ESC to continue.

**Appendix E: Proportion (and SD) of unrelated lures identified as old, as well as proportion (and SD) remembered, known, and guessed for initial recognition tests of Chapter 3. Data displayed by condition (strategy vs. no strategy) and lure type.**

| Condition | Lure type | Old | R | K | G |
|---|---|---|---|---|---|
| Strategy | Crit CM | .08(.28) | .03(.16) | .00(.00) | .05(.23) |
| | Crit consistent | .05(.16) | .01(.08) | .01(.08) | .03(.11) |
| | Control metaphor | .03(.11) | .00(.00) | .01(.08) | .01(.08) |
| | Control literal | .04(.14) | .01(.08) | .00(.00) | .03(.11) |
| | **Total[1]** | **.05(.11)** | **.01(.05)** | **.01(.03)** | **.03(.09)** |
| No strategy | Crit CM | .05(.23) | .03(.16) | .03(.16) | .00(.00) |
| | Crit consistent | .08(.19) | .04(.14) | .00(.00) | .04(.14) |
| | Control metaphor | .14(.25) | .04(.14) | .04(.14) | .05(.16) |
| | Control literal | .00(.00) | .00(.00) | .00(.00) | .00(.00) |
| | **Total[1]** | **.07(.10)** | **.03(.07)** | **.02(.05)** | **.02(.05)** |

Note: Some discrepancies due to rounding.
[1]Totals represent the average across the four lure types.

**Appendix F: Proportion (and SD) of unrelated lures identified as old, as well as proportion (and SD) remembered, known, and guessed for final recognition test of Chapter 3. Data displayed by repetition (previously tested vs. untested), condition (strategy vs. no strategy) and lure type.**

| Previously tested | Condition | Lure type | Old | R | K | G |
|---|---|---|---|---|---|---|
| Yes | Strategy | Crit CM | .19(.40) | .11(.32) | .03(.17) | .06(.23) |
| | | Crit consistent | .21(.30) | .11(.21) | .04(.18) | .06(.16) |
| | | Control metaphor | .21(.32) | .08(.22) | .08(.22) | .04(.14) |
| | | Control literal | .14(.31) | .04(.18) | .03(.12) | .07(.21) |
| | | **Total**[1] | **.19(.26)** | **.09(.19)** | **.05(.12)** | **.06(.10)** |
| | No strategy | Crit CM | .14(.35) | .03(.16) | .05(.23) | .05(.23) |
| | | Crit consistent | .26(.33) | .07(.17) | .09(.23) | .09(.23) |
| | | Control metaphor | .23(.32) | .04(.18) | .08(.19) | .11(.24) |
| | | Control literal | .14(.25) | .03(.11) | .04(.14) | .07(.17) |
| | | **Total**[1] | **.19(.21)** | **.04(.09)** | **.07(.11)** | **.08(.11)** |
| No | Strategy | Crit consistent | .03(.17) | .03(.17) | .00(.00) | .00(.00) |
| | | Control metaphor | .06(.23) | .03(.17) | .00(.00) | .03(.17) |
| | | Control literal | .01(.08) | .01(.08) | .00(.00) | .00(.00) |
| | | **Total**[1] | **.03(.11)** | **.02(.10)** | **.00(.00)** | **.01(.06)** |
| | No strategy | Crit consistent | .08(.28) | .00(.00) | .03(.16) | .05(.23) |
| | | Control metaphor | .11(.31) | .05(.23) | .00(.00) | .05(.23) |
| | | Control literal | .03(.11) | .00(.00) | .01(.08) | .01(.08) |
| | | **Total**[1] | **.07(.13)** | **.02(.08)** | **.01(.06)** | **.04(.11)** |

Note: Some discrepancies due to rounding.
[1]Totals represent the average across the four lure types.

**Appendix G: Proportion (and SD) of unrelated lures identified as old, as well as proportion (and SD) remembered, known, and guessed for initial recognition tests of Chapter 4. Data displayed by condition (full vs. divided attention) and lure type.**

| Condition | Lure type | Old | R | K | G |
|---|---|---|---|---|---|
| Full attention | Crit CM | .04(.20) | .02(.14) | .02(.14) | .00(.00) |
| | Crit consistent | .07(.20) | .01(.07) | .01(.07) | .05(.18) |
| | Control metaphor | .05(.15) | .01(.07) | .01(.07) | .03(.12) |
| | Control literal | .05(.15) | .02(.10) | .01(.07) | .02(.10) |
| | **Total**[1] | **.05(.09)** | **.02(.05)** | **.01(.05)** | **.03(.07)** |
| Divided attention | Crit CM | .20(.40) | .08(.27) | .02(.14) | .10(.30) |
| | Crit consistent | .30(.34) | .13(.24) | .06(.16) | .11(.21) |
| | Control metaphor | .31(.33) | .11(.23) | .06(.19) | .14(.25) |
| | Control literal | .24(.29) | .06(.16) | .08(.19) | .10(.20) |
| | **Total**[1] | **.26(.17)** | **.10(.14)** | **.06(.09)** | **.11(.13)** |

Note: Some discrepancies due to rounding.
[1]Totals represent the average across the four lure types.

**Appendix H: Proportion (and SD) of unrelated lures identified as old, as well as proportion (and SD) remembered, known, and guessed for final recognition test of Chapter 4. Data displayed by repetition (previously tested vs. untested), condition (full vs. divided attention) and lure type.**

| Previously tested | Condition | Lure type | Old | R | K | G |
|---|---|---|---|---|---|---|
| Yes | Full attention | Crit CM | .12(.33) | .06(.24) | .04(.20) | .02(.14) |
| | | Crit consistent | .26(.34) | .08(.21) | .08(.21) | .10(.25) |
| | | Control metaphor | .21(.32) | .05(.18) | .07(.18) | .09(.22) |
| | | Control literal | .22(.32) | .13(.24) | .03(.12) | .06(.16) |
| | | **Total[1]** | **.20(.22)** | **.08(.11)** | **.06(.09)** | **.07(.12)** |
| | Divided attention | Crit CM | .36(.48) | .14(.35) | .10(.30) | .12(.33) |
| | | Crit consistent | .35(.37) | .17(.28) | .05(.15) | .13(.26) |
| | | Control metaphor | .48(.38) | .20(.29) | .12(.28) | .16(.31) |
| | | Control literal | .40(.40) | .25(.35) | .07(.18) | .08(.21) |
| | | **Total[1]** | **.40(.25)** | **.19(.19)** | **.09(.15)** | **.12(.15)** |
| No | Full attention | Crit consistent | .04(.20) | .00(.00) | .04(.20) | .00(.00) |
| | | Control metaphor | .12(.33) | .02(.14) | .02(.14) | .08(.27) |
| | | Control literal | .02(.10) | .00(.00) | .01(.07) | .01(.07) |
| | | **Total[1]** | **.06(.13)** | **.01(.05)** | **.02(.08)** | **.03(.10)** |
| | Divided attention | Crit consistent | .12(.33) | .04(.20) | .02(.14) | .06(.24) |
| | | Control metaphor | .26(.44) | .04(.20) | .06(.24) | .16(.37) |
| | | Control literal | .16(.28) | .05(.15) | .02(.10) | .09(.22) |
| | | **Total[1]** | **.18(.24)** | **.04(.12)** | **.03(.11)** | **.10(.18)** |

Note: Some discrepancies due to rounding.
[1]Totals represent the average across the four lure types.

**Appendix I: Proportion (and SD) of unrelated lures identified as old, as well as proportion (and SD) remembered, known, and guessed for initial recognition tests of Chapter 5. Data displayed by attention at study (full vs. divided), attention at test (full vs. divided) and lure type.**

| Attention at Study | Attention at test | Lure type | Old | R | K | G |
|---|---|---|---|---|---|---|
| Full | Full | Crit CM | .18(.39) | .00(.00) | .08(.27) | .10(.31) |
| | | Crit consistent | .18(.27) | .03(.11) | .03(.11) | .13(.25) |
| | | Control metaphor | .09(.19) | .03(.11) | .04(.13) | .03(.11) |
| | | Control literal | .06(.20) | .01(.08) | .03(.11) | .03(.11) |
| | | **Total[1]** | **.13(.18)** | **.02(.04)** | **.04(.12)** | **.07(.11)** |
| | Divided | Crit CM | .08(.27) | .05(.22) | .03(.16) | .00(.00) |
| | | Crit consistent | .15(.26) | .08(.21) | .01(.08) | .06(.17) |
| | | Control metaphor | .10(.20) | .06(.17) | .04(.13) | .00(.00) |
| | | Control literal | .08(.21) | .05(.19) | .00(.00) | .03(.11) |
| | | **Total[1]** | **.10(.16)** | **.06(.14)** | **.02(.05)** | **.02(.05)** |
| Divided | Full | Crit CM | .15(.37) | .00(.00) | .08(.27) | .08(.27) |
| | | Crit consistent | .32(.35) | .03(.11) | .13(.25) | .17(.29) |
| | | Control metaphor | .32(.39) | .01(.08) | .10(.20) | .21(.30) |
| | | Control literal | .17(.26) | .03(.11) | .05(.15) | .09(.19) |
| | | **Total[1]** | **.24(.20)** | **.02(.05)** | **.09(.12)** | **.13(.12)** |
| | Divided | Crit CM | .25(.44) | .03(.16) | .05(.22) | .18(.38) |
| | | Crit consistent | .33(.33) | .05(.15) | .13(.22) | .15(.26) |
| | | Control metaphor | .35(.36) | .05(.15) | .16(.29) | .14(.25) |
| | | Control literal | .24(.34) | .03(.11) | .09(.19) | .13(.22) |
| | | **Total[1]** | **.29(.25)** | **.04(.08)** | **.11(.14)** | **.15(.14)** |

Note: Some discrepancies due to rounding.
[1]Totals represent the average across the four lure types.

**Appendix J: Proportion (and SD) of unrelated lures identified as old, as well as proportion (and SD) remembered, known, and guessed for final recognition test of Chapter 5. Data displayed by repetition (previously tested vs. untested), attention at study (full vs. divided), attention at test (full vs. divided) and lure type.**

| Previously tested | Attention at Study | Attention at test | Lure type | Old | R | K | G |
|---|---|---|---|---|---|---|---|
| Yes | Full | Full | Crit CM | .49(.51) | .15(.37) | .13(.34) | .21(.41) |
| | | | Crit consistent | .38(.42) | .14(.28) | .14(.30) | .10(.20) |
| | | | Control metaphor | .33(.42) | .13(.27) | .13(.25) | .08(.18) |
| | | | Control literal | .36(.40) | .12(.24) | .09(.19) | .15(.28) |
| | | | **Total[1]** | **.39(.33)** | **.13(.24)** | **.12(.17)** | **.13(.15)** |
| | | Divided | Crit CM | .38(.49) | .15(.36) | .03(.16) | .20(.41) |
| | | | Crit consistent | .35(.38) | .11(.27) | .09(.22) | .15(.28) |
| | | | Control metaphor | .36(.36) | .14(.28) | .10(.20) | .13(.25) |
| | | | Control literal | .25(.36) | .09(.25) | .09(.22) | .08(.21) |
| | | | **Total[1]** | **.33(.27)** | **.12(.20)** | **.08(.14)** | **.14(.16)** |
| | Divided | Full | Crit CM | .54(.51) | .18(.39) | .10(.31) | .26(.44) |
| | | | Crit consistent | .58(.34) | .14(.26) | .13(.27) | .31(.32) |
| | | | Control metaphor | .50(.38) | .09(.19) | .19(.32) | .22(.30) |
| | | | Control literal | .33(.37) | .08(.18) | .09(.19) | .17(.26) |
| | | | **Total[1]** | **.49(.27)** | **.12(.20)** | **.13(.14)** | **.24(.22)** |
| | | Divided | Crit CM | .53(.51) | .15(.36) | .03(.16) | .35(.48) |
| | | | Crit consistent | .45(.39) | .16(.29) | .10(.23) | .19(.27) |
| | | | Control metaphor | .46(.43) | .15(.28) | .14(.28) | .18(.31) |
| | | | Control literal | .45(.39) | .11(.29) | .14(.28) | .20(.35) |
| | | | **Total[1]** | **.47(.31)** | **.14(.19)** | **.10(.13)** | **.23(.23)** |
| No | Full | Full | Crit consistent | .05(.22) | .03(.16) | .00(.00) | .03(.16) |
| | | | Control metaphor | .10(.31) | .03(.16) | .00(.00) | .08(.27) |
| | | | Control literal | .08(.22) | .03(.11) | .00(.00) | .05(.15) |
| | | | **Total[1]** | **.08(.20)** | **.03(.11)** | **.00(.00)** | **.05(.13)** |
| | | Divided | Crit consistent | .20(.41) | .08(.27) | .00(.00) | .13(.33) |
| | | | Control metaphor | .08(.27) | .03(.16) | .00(.00) | .05(.22) |
| | | | Control literal | .10(.26) | .03(.11) | .01(.08) | .06(.20) |
| | | | **Total[1]** | **.13(.19)** | **.04(.12)** | **.00(.03)** | **.08(.15)** |
| | Divided | Full | Crit consistent | .15(.37) | .03(.16) | .05(.22) | .08(.27) |
| | | | Control metaphor | .28(.46) | .03(.16) | .05(.22) | .21(.41) |
| | | | Control literal | .13(.30) | .01(.08) | .01(.08) | .10(.26) |
| | | | **Total[1]** | **.19(.26)** | **.02(.08)** | **.04(.10)** | **.13(.23)** |
| | | Divided | Crit consistent | .25(.44) | .08(.27) | .08(.27) | .10(.30) |
| | | | Control metaphor | .30(.46) | .05(.22) | .10(.30) | .15(.36) |
| | | | Control literal | .19(.33) | .08(.24) | .03(.11) | .09(.19) |
| | | | **Total[1]** | **.25(.31)** | **.07(.20)** | **.07(.17)** | **.11(.22)** |

Note: Some discrepancies due to rounding.
[1]Totals represent the average across the four lure types.

## Appendix K: High-threshold correction procedure

To apply the "high-threshold correction," the false recognition proportion when the lure is unrelated to the study list is subtracted from the false recognition proportion when it is related (Gallo & Roediger, 2002; Schacter et al., 1996; Seamon et al., 1998). In Chapter 5, the unrelated critical consistent lures were falsely recognized more often than the unrelated control literal lures for both the initial recognition tests (.244 vs. .136) and for the repeated lures on the final recognition test (.440 vs. .348). However, there were no reliable interactions between lure type and attention at either study or test, so the overall unrelated false recognition proportions averaged across all conditions were used as the subtraction values. These values were subtracted from the main false recognition proportions (i.e., when these lures were related to the study list) for each participant. For example, on the initial recognition tests, one participant had a false recognition proportion of .5 for the critical consistent lures and .25 for the control literal lures when these lures were related to the study lists. After the adjustment was applied, the proportions would be .256 and .114 respectively. The mean adjusted values for the initial recognition tests are displayed in Table K-1 and for the final recognition test are displayed in Table K-2.

**Table K-1.** Comparison of adjusted false recognition proportions for critical consistent lures and control literal lures. Data are from the initial recognition tests in Chapter 5. Unadjusted proportions, the adjustment calculation, and the final adjusted values are presented.

| Attention at study | Attention at test | Lure type | Unadjusted false recognition | Adjustment | Adjusted false recognition[1] |
|---|---|---|---|---|---|
| Full | Full | Crit consistent | .29(.24) | – .244 | .05(.24) |
| | | Cont literal | .18(.20) | – .136 | .05(.20) |
| | Divided | Crit consistent | .33(.23) | – .244 | .09(.23) |
| | | Cont literal | .16(.19) | – .136 | .03(.19) |
| Divided | Full | Crit consistent | .54(.22) | – .244 | .29(.22) |
| | | Cont literal | .36(.26) | – .136 | .22(.26) |
| | Divided | Crit consistent | .52(.26) | – .244 | .27(.26) |
| | | Cont literal | .36(.24) | – .136 | .23(.24) |

[1]Some discrepancies due to rounding.

**Table K-2.** Comparison of adjusted false recognition proportions for repeated critical consistent lures and repeated control literal lures. Data are from the final recognition test in Chapter 5. Unadjusted proportions, the adjustment calculation, and the final adjusted values are presented.

| Attention at study | Attention at test | Lure type (Repeated) | Unadjusted false recognition | Adjustment | Adjusted false recognition[1] |
|---|---|---|---|---|---|
| Full | Full | Crit consistent | .61(.28) | – .440 | .17(.28) |
| | | Cont literal | .36(.31) | – .348 | .01(.31) |
| | Divided | Crit consistent | .60(.25) | – .440 | .16(.25) |
| | | Cont literal | .29(.26) | – .348 | -.05(.26) |
| Divided | Full | Crit consistent | .73(.20) | – .440 | .29(.20) |
| | | Cont literal | .47(.23) | – .348 | .13(.23) |
| | Divided | Crit consistent | .69(.23) | – .440 | .25(.23) |
| | | Cont literal | .50(.27) | – .348 | .15(.27) |

[1]Some discrepancies due to rounding.

# Appendix L: Ethics approval notices

Western Research

**Date:** 21 December 2017

**To:** Dr. Albert Katz

**Project ID:** 110444

**Study Title:** The Effect of Language on Memory

**Application Type:** NMREB Initial Application

**Review Type:** Delegated

**Full Board Reporting Date:** January 12 2018

**Date Approval Issued:** 21/Dec/2017

**REB Approval Expiry Date:** 21/Dec/2018

---

Dear Dr. Albert Katz

The Western University Non-Medical Research Ethics Board (NMREB) has reviewed and approved the WREM application form for the above mentioned study, as of the date noted above. NMREB approval for this study remains valid until the expiry date noted above, conditional to timely submission and acceptance of NMREB Continuing Ethics Review.

This research study is to be conducted by the investigator noted above. All other required institutional approvals must also be obtained prior to the conduct of the study.

**Documents Approved:**

| Document Name | Document Type | Document Date | Document Version |
|---|---|---|---|
| Debriefing form for the effect of langauge on memory clean -- clean copy | Debriefing document | 06/Oct/2017 | 1 |
| LOI -- The Effect of Language on Memory -- Clean copy | Written Consent/Assent | 18/Nov/2017 | 1 |
| Recruitment Poster for SONA -- The Effect of Langauge on Memory -- Clean copy | Recruitment Materials | 18/Nov/2017 | 2 |
| Stimuli Metaphor and Memory | Other Data Collection Instruments | 06/Oct/2017 | 1 |

No deviations from, or changes to the protocol should be initiated without prior written approval from the NMREB, except when necessary to eliminate immediate hazard(s) to study participants or when the change(s) involves only administrative or logistical aspects of the trial.

The Western University NMREB operates in compliance with the Tri-Council Policy Statement Ethical Conduct for Research Involving Humans (TCPS2), the Ontario Personal Health Information Protection Act (PHIPA, 2004), and the applicable laws and regulations of Ontario. Members of the NMREB who are named as Investigators in research studies do not participate in discussions related to, nor vote on such studies when they are presented to the REB. The NMREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000941.

Please do not hesitate to contact us if you have any questions.

Sincerely,

[redacted]

***Note: This correspondence includes an electronic signature (validation and approval via an online system that is compliant with all regulations).***

Western Research

**Date:** 3 April 2019

**To:** Dr. Albert Katz

**Project ID:** 110444

**Study Title:** The Effect of Language on Memory

**Application Type:** NMREB Amendment Form

**Review Type:** Delegated

**Full Board Reporting Date:** May 3 2019

**Date Approval Issued:** 03/Apr/2019

**REB Approval Expiry Date:** 21/Dec/2019

---

Dear Dr. Albert Katz,

The Western University Non-Medical Research Ethics Board (NMREB) has reviewed and approved the WREM application form for the amendment, as of the date noted above.

**Documents Approved:**

| Document Name | Document Type | Document Date | Document Version |
|---|---|---|---|
| Language and memory email recuritment script 2019-03-29 | Recruitment Materials | 29/Mar/2019 | 1 |
| Language and memory Recruitment Poster | Recruiting Advertisements | 29/Mar/2019 | 1 |
| Language and memory website ad 2019-03-29 | Recruiting Advertisements | 29/Mar/2019 | 1 |
| LOI -- The Effect of Language on Memory 2019-03-29 PAID PARTICIPANTS | Written Consent/Assent | 29/Mar/2019 | 1 |
| LOI -- The Effect of Language on Memory 2019-03-29 SONA PARTICIPANTS | Written Consent/Assent | 29/Mar/2019 | 1 |
| Recruitment Poster for SONA -- The Effect of Langauge on Memory 2019-03-29 | Recruitment Materials | 29/Mar/2019 | 1 |

REB members involved in the research project do not participate in the review, discussion or decision.

The Western University NMREB operates in compliance with the Tri-Council Policy Statement Ethical Conduct for Research Involving Humans (TCPS2), the Ontario Personal Health Information Protection Act (PHIPA, 2004), and the applicable laws and regulations of Ontario. Members of the NMREB who are named as Investigators in research studies do not participate in discussions related to, nor vote on such studies when they are presented to the REB. The NMREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000941.

Please do not hesitate to contact us if you have any questions.

Sincerely,

[redacted]

***Note: This correspondence includes an electronic signature (validation and approval via an online system that is compliant with all regulations).***

# Curriculum Vitae

| | |
|---|---|
| **Name:** | Nick Reid |
| **Post-secondary Education and Degrees:** | University of New Brunswick<br>Saint John, New Brunswick, Canada<br>2008-2014 B.A. |
| | The University of Western Ontario<br>London, Ontario, Canada<br>2014-2016 M.Sc. |
| | The University of Western Ontario<br>London, Ontario, Canada<br>2016-2020 Ph.D. |
| **Honours and Awards:** | Ontario Graduate Scholarship<br>2018-2019, 2019-2020 |
| | Social Science and Humanities Research Council (SSHRC)<br>Canadian Graduate Scholarship – Masters<br>2014-2015 |
| **Related Work Experience** | Teaching Assistant<br>The University of Western Ontario<br>2014-2019 |

**Peer-reviewed articles:**

**Reid, J. N.,** Al-Azary, H., & Katz, A. N. (in press). Metaphors: Where the neighborhood in which one resides interacts with (interpretive) diversity. In *Proceedings of the 42nd Annual Meeting of the Cognitive Science Society*.

Katz, A. & **Reid, J. N.** (2020). Tests of Conceptual Metaphor Theory with episodic memory tests. *Cognitive Semantics, 6*(1), 56-82. https://doi.org/10.1163/23526416-00601003

**Reid, J. N.,** & Katz, A. N. (2018a). Something false about conceptual metaphors. *Metaphor and Symbol, 33*(1), 36-47. https://doi.org/10.1080/10926488.2018.1407994

**Reid, J. N.,** & Katz, A. N. (2018b). Vector space applications in metaphor comprehension. *Metaphor and Symbol, 33*(4), 280-294. https://doi.org/10.1080/10926488.2018.1549840

**Articles under review:**

**Reid, J. N.,** & Katz, A. N. (2020). *The RK processor: A program for analysing metaphor and word feature listing data.* Manuscript being revised for resubmission.

Yang, H., **Reid, J. N.,** Katz, A. N., & Li, D. (2020). *The embodiment of power as forward/backward movement in Chinese and English Speakers.* Manuscript submitted for publication

**Conference presentations:**

(* = presenting author)

*__Reid, J. N.,__ & Katz, A. N. (2020, June). *Conceptual metaphor activation is automatic.* Talk presented at the 13th Researching and Applying Metaphor Conference. Virtual meeting. Note: Winner of the "Best PhD conference paper" prize.

*__Reid, J. N.,__ & Katz, A. N. (2019, November). *Conceptual metaphors influence memory automatically: Evidence from a divided attention false memory task.* Poster presented at the 60th annual meeting of the Psychonomic Society, Montréal, Québec, Canada.

*Yang, H., **Reid, J. N.,** & Katz, A. N. (2019, November). *Social power and approach behaviour in English vs. Mandarin speakers.* Poster presented at the 60th annual meeting of the Psychonomic Society, Montréal, Québec, Canada.

*__Reid, J. N.,__ & Katz, A. N. (2019, June). *Conceptual metaphors influence cognition, but is the effect conscious?* Talk presented at the 23rd annual meeting of the Association for the Scientific Study of Consciousness, London, Ontario, Canada.

*__Reid, J. N.,__ & Katz, A. N. (2019, June). *The conceptual metaphor false memory effect.* Talk presented at the 29th annual meeting of the Canadian Society for Brain, Behaviour and Cognitive Science, Waterloo, Ontario, Canada.

*__Reid, J. N.,__ & Katz, A. N. (2019, February). *Conceptual metaphors induce false memories: Evidence from a remember/know recognition test.* Poster presented at the 48th Lake Ontario Visionary Establishment Conference, Niagara Falls, Ontario, Canada.

*__Reid, J. N.,__ & Katz, A. N. (2018, June). *Testing Conceptual Metaphor Theory using memory methodologies.* Talk presented at the 12th Researching and Applying Metaphor Conference, Hong Kong.

Katz, A. N., & ***Reid, J. N.*** (2017, November). *An empirical test of Conceptual Metaphor Theory.* Poster presented at the 58th annual meeting of the Psychonomic Society, Vancouver, BC, Canada.

Katz, A. N., & ***Reid, J. N.*** (2016, November). *Perception of spatial distance between two people after reading metaphor.* Poster presented at the 57th annual meeting of the Psychonomic Society, Boston, MA.