EFFECT OF LOW- AND HIGH-AROUSAL MOOD ON TRUE AND FALSE MEMORY FOR NEUTRAL WORDS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Sarah Bookbinder

December 2017

EFFECT OF LOW- AND HIGH-AROUSAL MOOD ON TRUE AND FALSE MEMORY FOR NEUTRAL WORDS

Sarah Bookbinder, Ph.D.

Cornell University 2017

This dissertation examined the contributions of mood valence and arousal to false memory (FM) for neutral DRM word lists. There has been a lively debate in the literature regarding whether valence or arousal is the primary cause of mood-dependent FM. Because much past research has not effectively disentangled valence from arousal, the present studies used a new mood manipulation that allowed valence and arousal to be controlled and their associated retrieval processes to be measured. In Experiment 1, low-arousal negative moods reduced false memory compared to positive and neutral moods by increasing verbatim memory traces, confirming the prediction that valence can influence false memory independent of arousal. In Experiment 2, low-arousal negative moods reduced false memory for both list lengths, but the processes associated with false memory for short lists were not the same as the processes associated with false memory for long lists. High arousal negative moods in Experiment 3 now increased false memory compared to neutral moods—but still reduced false memory compared to positive moods—which suggested that high arousal increased false memory compared to low arousal. Experiment 4 provided a direct test of arousal's effect on false memory and confirmed that higher levels of arousal increase false memory via strong gist traces, regardless of valence. The

results were discussed in support of fuzzy-trace theory—but not affect-as-information theory—and practical applications to the law were noted.

BIOGRAPHICAL SKETCH

Sarah Bookbinder earned her B.A. in Psychology from Hamilton College in 2010 and earned her M.A. in Developmental Psychology from Cornell University in 2015. She grew up in New Canaan, CT and has resided in Ithaca, NY while earning her PhD in Developmental Psychology from Cornell University's Department of Human Development. Her current research interests lie in the field of cognitive psychology and its relationship to the law, with primary interest in the way that emotions affect memory processes and their implications for creating false memories in legal settings.

ACKNOWLEDGMENTS

I have several people to thank for the creation of this work and the support provided to me along the way. First, I thank Charles Brainerd for his mentorship, both directing me and allowing me to take the lead. I also thank Melissa Ferguson, for advising me through this process from the beginning and helping me expand my focus, and Steve Ceci, for providing guidance through the most helpful of questions and comments. I must also thank the hardworking research assistants in the lab, without whom these data could not exist. I also must acknowledge the advice, support, and sometimes commiseration of the women in the department who became not just colleagues but lifelong friends. Finally, I give huge thanks to my parents, my sister, and Frank, whose patience and support have been endless.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AAI	Affect-as-information
CDs	Critical distractors
CR	Conjoint recognition
DRM	Deese/Roediger/McDermott paradigm
EEM	Emotional enhancement of memory
FM	False memory
FTT	Fuzzy-trace theory
G	Gist; gist memory instruction
MIP	Mood induction procedure
PNE	Paradoxical negative emotion
RDs	Related distractors
TM	True memory
URDs	Unrelated distractors
V	Verbatim; verbatim memory instruction
VG	Verbatim-plus-gist

CHAPTER 1

INTRODUCTION

The most memorable moments in our lives are often those most charged with emotion. Memory accuracy—and the feeling of accuracy, which is not the same thing as real accuracy—for these events may depend on the nature of that emotion. Even during casual conversation people exhibit belief in a relationship between our emotions and our memories. One common belief is that the details of highly positive events become blurred. We might discuss a fond memory of an event such as a wedding by recalling how much fun we had but not being able to account for what we did during each hour. Another belief is that we are highly accurate at remembering the details of highly negative events. After witnessing a disaster or terrorist attack, people often report that they can recall every detail with striking clarity, as if they can re-watch a movie of the event in their mind's eye.

The latter type of memory—for negative events—has been the subject of much research because of its many real-world applications and because of the longstanding belief that memory for negative events is particularly vivid and accurate. Much focus has been devoted to flashbulb memory, which is the idea that highly emotional, negative events ensure memory accuracy (Brown & Kulik, 1977). However, research testing this flashbulb memory effect has shown that although *feelings* of accuracy and vividness were high for memories of events such as the 9/11 terrorist attacks, actual accuracy was not (Talarico & Rubin, 2003). Other evidence similarly points to the idea that emotional (particularly negative) experiences may lead to memory inaccuracies. For example, during police interviews witnesses may be especially prone to misremembering the details of crimes (Brainerd & Reyna, 2005). In some cases, negative emotion has been associated with false memories for events that never even occurred (and are highly unlikely to occur, such as an alien abduction) (e.g., Shaw & Porter, 2015; Spanos, 1996).

The association between false memory and negative emotion has clear practical implications because the situations in which consequences of false memories are the most serious are also highly emotional situations. Police interviews, eyewitness testimony, and recounting of abuse during therapy are three notable examples. In these situations, and in the types of situations assessed in the research mentioned above, emotion plays a role in two ways. First, the event that occurs (or the materials used in the laboratory) may be emotional itself (e.g., a crime, disaster, or emotional words or pictures). I refer to this aspect of emotion as *emotional content*. Second, people's moods while the events are experienced may also be emotional (e.g., fearful, angry, or happy). I refer to this aspect of emotion as *mood*, generally, but it has also been referred to as *emotional context* in the literature (see Bookbinder & Brainerd, 2016). Thus, emotion can figure into memory situations in two different ways, and the qualities of emotional content and mood may not necessarily match or remain stable over time. For example, someone may be fearful while witnessing a crime, but after repeated police interviews may become irritated when repeatedly recounting the same details, although the negative evaluation of the crime remains the same.

In order to determine precisely how emotion affects false memory, then, it is necessary to investigate the effects of mood and content separately, since there is evidence that their effects on memory are not the same (Bookbinder & Brainerd, 2016). Research using a variety of materials has demonstrated fairly consistently that negative emotional content elevates false memory compared to positive content (e.g., Bookbinder & Brainerd, 2017; Brainerd, Stein, Silveira, Rohenkohl, & Reyna, 2008). Conversely, research using experimentally-induced moods has shown that positive moods elevate false memory compared to negative moods (e.g., Storbeck, 2013; Storbeck & Clore, 2005), not to mention naturally-occurring moods, such as those related to psychopathy, which have an entirely different relationship with memory (see Otgaar, Muris, Howe, & Merckelbach, 2017). However, in both content and mood research the results have been difficult to interpret because of the conflation of two dimensions of emotion. As will be discussed in the upcoming theoretical section, emotion is commonly thought to vary along the dimensions of valence and arousal. In some emotion research, however, valence and arousal have been confounded, prohibiting researchers from drawing conclusions about whether valence, arousal, or both affect memory. In the study of emotional content, this problem has the beginning of a solution thanks to research that controls one dimension while manipulating the other or factorially manipulates the two. These data suggest that valence has a fairly consistent effect on false memory (for a review see Bookbinder & Brainerd, 2016). In the study of mood, however, precise control of valence and arousal is more difficult to achieve. Creating mood inductions that induce reliable mood changes in subjects is difficult in the laboratory to begin with, and it is particularly difficult to induce reliable changes in mood that are strong enough to affect memory while keeping arousal levels low; when arousal is kept low, levels of valence that are reliably different are difficult to achieve. On the other hand, when creating levels of positive and negative valence that are not only different from each other but also different from neutral moods, arousal levels are apt to become higher for positive and negative moods compared to neutral moods. Thus, much research has focused on comparing highly arousing positive and negative moods to neutral moods, which leaves open the question of what combination of valence and arousal is necessary and sufficient to affect false memory.

The present research addresses this open question by creating a mood induction procedure (MIP) that allows valence and arousal to be separated and by measuring the effect of this MIP on true and false memory. Furthermore, this series of experiments is the first to additionally measure the retrieval processes associated with true and false memory, for subjects experiencing a variety of moods and for several other factors that may influence the mood-memory relationship. In the following sections of the first chapter of this dissertation I will discuss the theoretical framework that led to my predictions about mood effects on false memory. Chapters 2 through 6 detail the pilot tests and the four studies comprising this line of research. Finally, Chapter 7 provides a general discussion of these results and their implications. To begin the theoretical background for this research I discuss theories of emotion and how they have been applied to theories of memory.

Theoretical Background: False Memory

It is necessary to first clarify the type of false memory that this research deals with and the theoretical explanations for it. I then discuss how false memory may be measured and manipulated, focusing on emotion as one factor that may influence false memory.

When I mention that I study false memory to a friend outside my field, they often show keen interest as they imagine me convincing subjects of absurd past experiences such as going swimming with the pope or owning a pet monkey. In reality, the false memories that researchers typically measure are much more mundane. They are not usually events that subjects have never experienced (e.g., owning a monkey), but rather, they are false in that they have never been experienced in a particular context (e.g., falsely remembering that you had a pirate-themed birthday party when really it was your sibling's party). (Although, as mentioned earlier, some interesting research has shown that it is possible to implant memories of absurd, unlikely events). The typical sort of false memories is also not usually related to highly unfamiliar events (e.g., the pope [for most people]) but rather are related to familiar events or items (e.g., birthday parties). Finally, false memories are usually congruous with the general meaning content, or gist, of an event that did happen (e.g., remembering that you ate cake at the birthday party when you really had cupcakes, rather than remembering that you had gruel at a birthday party, which is not consistent with the gist of birthday parties). In this way, false memories are not fantastical, random memories that have no grounds in reality but instead are memory distortions where the meaning of a real event is preserved. There are two main types of false memory that fit these criteria.

Types of False Memory

Spontaneous false memories occur as a result of trying to remember events that fit with a particular gist of experienced events. For example, if you go to the zoo and see a lion, a tiger, and a leopard, when you get home you might try to remember the animals you saw and falsely remember that you saw a cheetah as well. This error occurs because *cheetah* fits with the gist of what you experienced: big cats that are typically seen at the zoo. Research on spontaneous false memories might use pictures, word lists, narratives, or even videos. It can be seen how the zoo scenario could be adopted for any of those presentation modalities. Subjects' memories for those presented events are typically measured using either recall or recognition tests. On recall tests subjects are simply asked to list everything they remember experiencing (e.g., tiger, cheetah [false alarm], lion). On recognition tests, subjects are presented with targets (presented items to measure true memory; lion), related distractors (RDs; unpresented items to measure false memory; cheetah), and unrelated distractors (URDs; unpresented unrelated items to measure response bias; goldfish) and respond “yes” or “no” to whether that item was previously presented. In addition to simply responding to whether an item has been presented or not, sometimes subjects are asked to rate how confident they are in their response, or to decide whether they “remember” that an item was presented earlier because of conscious recollection of details, or whether they simply “know” that it was presented earlier because of a more general sense of familiarity (remember/know procedure; Tulving, 1985).

The Deese/Roediger/McDermott paradigm (DRM; Deese, 1959; Roediger & McDermott, 1995); is the most commonly used method of studying spontaneous false memories and is the method used in the present research. Subjects listen to (or read) lists of words (e.g., table, sit, desk, bench) that are all related to a single unpresented critical distractor word (CD; chair). They then complete recall or recognition tests in which the unpresented critical distractors are the false memory items. Other related words (RDs) may serve as additional false memory items. These lists reliably produce high levels of false memory and can easily be adapted to study emotional effects. Researchers have created variations of the standard neutral DRM lists in which the CDs are emotional (e.g., Budson et al., 2006) and researchers have also implemented mood induction procedures prior to presentation of neutral study lists (e.g., Corson & Verrier, 2007). The latter method is the one adopted in the present research. In a third type of design, researchers have used mood induction techniques along with emotional rather than neutral lists, in order to test mood congruency effects on false memory (e.g., Ruci, Tomes, & Zelenski, 2009).

The other main type of false memory does not factor into the present research, but I discuss it here briefly for completeness and to provide contrast. Imagine that after you went to the zoo you were recounting your experience to a friend and she asks whether you saw a cheetah (remember that you did not). “Are you *sure* you didn’t see a cheetah? They always have cheetahs at the zoo,” she insists. In this case you might be quite apt to agree and say that yes, you saw a cheetah even though you did not, because what is being suggested to you is consistent with the gist of your zoo experience. If your friend asked if you saw a fireman at the zoo, however, you would be less likely to misremember and say yes to that. Whereas spontaneous false memories can be fully attributable to endogenous distortion, here any false memories that occur are additionally influenced by external distortion (your friend’s question, social pressure to agree with your friend, etc.) as well. This fact introduces another element into the emotion-false memory equation, but one that is not an issue in the present spontaneous false memory design.

Factors Affecting the DRM Illusion

In addition to emotion, which will be discussed as a false memory factor in the next section, several factors have been manipulated within the DRM paradigm to determine the conditions under which true memory and false memory will increase or decrease. In one of the earliest studies of this nature, Roediger, Watson, McDermott, and Gallo (2001) identified seven factors that can account for variance in false recall rates, with the most variance accounted for by backward associative strength (BAS). BAS is correlated with both false recall and false recognition and refers to the association from the targets to the CD.

Several other characteristics of study lists as well as the way the lists are tested influence DRM false memory. Increasing the number of targets, or the list length, increases false recall and recognition, but reduces true memory (Robinson & Roediger, 1997; Sugrue & Hayne, 2006).

When lists of 7 and 14 targets were compared, false recall was higher for lists of 14 targets. DRM lists that are presented visually reduce false memory compared to lists that are presented auditorily (Gallo, McDermott, & Percer, 2001; Smith & Hunt, 1998), and the length of time that each word is presented, or the time between words, matters as well. Increasing presentation duration increases false memory up to a point, but after about 5000ms false memory rates begin to decrease (Gallo & Roediger, 2002; McDermott & Watson, 2001). Finally, how subjects encode study lists may influence their later memory as well. Deeper processing, which encourages subjects to consider words' meaning, increases false memory compared to shallow processing, which encourages subjects to focus on visual or phonemic characteristics (Rhodes & Anastasi, 2000; Thapar & McDermott, 2001).

Two other factors relate to how memory is tested rather than how lists are presented. When subjects' memories are tested more than once, earlier memory tests may influence their responding on later memory tests. For example, half of the subjects in Roediger and McDermott's (1995) Experiment 2 completed recall tests after each list, while the other half did math problems after each list. When a final recognition test was given after all lists had been presented, false recognition was higher for subjects who had done recall tests earlier as opposed to math problems. Similarly, when initial recall tests were administered, later false recall was higher compared to subjects who did not do earlier recall tests, and true recall was higher as well (McDermott, 2006). Furthermore, remember (as opposed to know) judgments were more common, for both targets and CDs, after repeated testing. Thus, although repeated testing has beneficial effects on true memory, it is detrimental to false memory to the extent that falsely remembered words are specifically remembered, rather than generally known.

Finally, the amount of time between studying targets and the memory test, known as retention interval, may also affect false memory, although there has historically been some inconsistency on that topic. On one hand, McDermott (1996) found that false memory remains stable over a two-day delay, whereas on the other hand, Lampinen and Schwartz (2000) found that false memory rates declined over two days. What is relatively consistent, however, is the finding that true memory declines rapidly over time, and that that decline is faster than the decline of false memories (Seamon, Luo, Shulman, Toner, & Caglar, 2002; Thapar & McDermott, 2001; Toglia, Neuschatz, & Goodwin, 1999).

These variations on the DRM illusion have informed the several theoretical explanations for the illusion, which will be discussed next. Emotion, of course, is another potential influencer of DRM false memory, which I discuss in detail in the next section. Importantly, the way that emotion-memory effects depend on the above factors may provide insight into the process basis for emotional effects.

Theories of False Memory

Early theorists of false memory proposed that true and false memory relied on a single process, resulting in predictions that true and false memory would be correlated and would be affected in the same way by any one manipulation. One of the first of such theories, implicit associative response theory, explained two responses made following to-be-remembered words: a representational response, directly capturing the presented words themselves, and an implicit associative response, capturing unpresented words associated with studied words (Underwood, 1965). When implicit associative responses are confused with representational responses false memory occurs. Furthermore, the more often an implicit associative response occurs, the more likely false memories become, which explains why DRM lists with more associated targets results in higher levels of false memory (even when the total number of list words is held constant, see Robinson & Roediger, 1997). However, implicit associative response theory was not able to account for false memories' persistence over time, because implicit associative activation decays quickly (Anderson, 1983).

Later, schema theory relied on the idea that target events are encoded into schemata, and retrieval from those schemata results in false memories (Alba & Hasher, 1983). Similarly, the source-monitoring framework (Johnson, Hashtroudi, & Lindsay, 1993) proposed that true memories included source attributions about the context of events; relying on source attributions could lead to false memories as well. In these original conceptions, increasing subjects' understanding of a schema or source would not only enhance their true memory but would increase the likelihood of memory distortion as well. That is, by these accounts true and false memory should be correlated, and any manipulation that increases true memory should increase false memory as well.

However, early research testing the two predictions of one-process theories did not reveal support for either one. True and false memory are usually uncorrelated (e.g., Reyna & Kiernan, 1994) and some manipulations can affect true and false memory in opposite ways, or affect one type of memory but not the other (e.g., Schacter, Israel, & Racine, 1999). For example, the effect of list length described above—increasing the number of presented targets from 7 to 14—increased false memory while reducing true memory. Since one-process accounts could not explain these results, opponent-process theories emerged, providing a better account of unparallel true and false memory effects.

Opponent-process theories draw on the concepts of dual-process theory, which describes two retrieval processes that contribute to (correctly) remembering an item. Recollection involves the conscious reinstatement of details of the item's presentation and familiarity involves strong feelings that the item was presented without awareness of its details. These two retrieval processes contribute to falsely remembering an item as well, but their contributions are opposite in this case. Recollection suppresses false memories because it consciously reinstates details of the items' presentation, which enables the rejection of similar items whose details do not match presented details. Familiarity, on the other hand, supports false memories because the details that related/critical distractor items share with target items produce strong feelings that they were presented.

These two processes can be dissociated using the remember/know paradigm described above as well as other methods such as process dissociation (Jacoby, 1991) and item versus associative recognition tasks (Yonelinas, 1997). The results from these various methods provided evidence for separate neural regions controlling the two processes, that these processes are independent at retrieval, and that they exhibit different processing speeds (for a review see Yonelinas, 2002). On the other hand, others have found overlap in the brain regions associated with TM and FM (Dennis, Bowman, & Vandekar, 2012).

Several opponent-process theories provide specific predictions about true and false memory effects and the processes driving them. Spreading activation theory (McDermott & Watson, 2001; Roediger, Balota, & Watson, 2001) was originally developed to explain semantic processing and was later applied to the DRM illusion. By this account, words and concepts are semantically organized in mental storage, so that processing one word activates its mental node, which activates surrounding, related nodes that correspond to associated words. False memories occur when activation of nodes of studied words causes activation of nodes of semantically related CDs. False memories can be prevented when strategic processing allows subjects to differentiate experienced events from activated events. Two theories grew out of this theoretical framework: associative activation theory and activation monitoring theory.

Activation monitoring theory (AMT; Roediger, Watson, McDermott, & Gallo, 2001) expanded upon spreading activation theory by conceptualizing a broader type of activation that is not limited to semantic associations, but can include phonological, orthographic, or other associations. Study lists activate concepts related via semantic or other connections, which may activate other related concepts, and thereby activate CDs. Again, on the memory test, subjects may be unable to differentiate studied words from words that were activated during the study phase (even though they were not presented). Lists that have high levels of association to related concepts will thus produce high levels of false memory. This theory can explain why repeated testing increases false memory, because recalling CDs or seeing them on a recognition test strengthens activation and makes it more difficult to differentiate them from targets on a later memory test. It also provides an explanation for false memory reduction through the concept of monitoring. Differences between studied words that have actually been experienced and activated words that have not been experienced can be used to differentiate words during a memory test and can be used to correctly reject CDs. The stronger the activation of CDs, the more difficult it may be to monitor and reject them during a memory test. Monitoring may also occur during encoding, which explains why false memory declines at longer presentation durations, because subjects have more time to monitor. By this account, then, lists that produce high levels of activation, or low levels of monitoring, will produce high levels of false memory. However, AMT is specific to words, and cannot explain false memory's persistence over time compared to true memory.

The final opponent process theory I will discuss, which provides the motivation for the present research, can be applied to stimuli other than words and provides an explanation for the time span of true and false memory. This theory is different from activation theories because rather than explaining how targets may activate lexical representations of other words, this theory makes predictions regarding the type of memory traces that are stored for targets during item presentation and later retrieved during a memory test. According to fuzzy-trace theory (FTT; Reyna & Brainerd, 1995) verbatim traces store items' surface form, are the basis for recollection, and support true memory but suppress false memory. Gist traces, on the other hand, store items' meaning, are the basis for familiarity, and support both true and false memory.

According to AMT, stronger word associations elevate false memories, whereas according to FTT, stronger gist (a concept which can be applied to pictures and other stimuli aside from words) strengthens false memories. A comparison of the predictions of FTT and associative theories (Dewhurst, Pursglove, & Lewis, 2007) revealed that when gist was strengthened without strengthening word associations, by presenting targets within a narrative, false memories increased. Furthermore, false memory effects can be found for pictures as well as words, which cannot easily be explained by AMT (e.g., Bookbinder & Brainerd, 2017), and FTT predicts that verbatim traces fade more rapidly than gist traces because surface details become less clear over time.

Another notable finding in the false memory literature is that false memories are often accompanied by remember judgments rather than know judgments, signifying strong feelings of remembering that they were previously presented even though they were not. FTT explains this phenomenon via the concept of phantom recollection: When gist traces are sufficiently strong, they can create vivid mental reinstatement of a CD's prior occurrence despite the fact that it was not ever experienced. Phantom recollection may thus explain why strong feelings of remembering or high confidence may accompany false memories and cannot be relied upon as a measure of accuracy. Associative theories do not provide a similar explanation for remember judgments of false memories.

Concepts of FTT may be applied to explain the way that emotion affects false memory, specifically, as well. Emotion—in the form of content or mood—could potentially affect verbatim traces or gist traces. A review of the existing emotion-false memory literature points to the general conclusion that emotional content affects gist traces whereas emotional context has a greater effect on verbatim traces (Bookbinder & Brainerd, 2016). Emotion's effect on false memory can be further defined at the level of specific retrieval processes relying on either verbatim or gist memory traces. Existing research has begun to investigate the question of retrieval processes in false memory for emotional words (Brainerd, Stein, Silveira, Rohenkohl, & Reyna, 2008) and emotional pictures (Bookbinder & Brainerd, 2017). The results of those studies will be reviewed in the next section. The present research is the first to measure retrieval processes involved in false memory related to emotional mood.

Conjoint Recognition

The methodology designed to measure verbatim and gist retrieval processes is known as conjoint recognition (CR; Brainerd, Wright, Reyna, & Mojardin, 2001). This procedure can be adapted in any memory paradigm by altering the standard recall or recognition test. Instead of asking subjects "was this item presented?" to which they should answer "yes" or "no," a conjoint recognition test asks three questions: (1) was this item presented? (the same question asked on standard recognition tests), (2) was a similar item presented?, or (3) was either this item or a similar item presented? These questions are referred to as verbatim (V), gist (G), and verbatim-plus-gist (VG) questions and can be applied either within- or between-subjects. In order to measure underlying retrieval processes a multinomial model is applied to the resulting data. As displayed in Table 1.1, these retrieval processes relate to the processing of verbatim or gist traces and support either true or false memory.

Table 1.1

Parameters of the Conjoint Recognition Model

Parameter	Definition
	False Memory Processes
P_D	Phantom recollection (RD produces retrieval of gist trace of similar target and false acceptance of RD)
R_D	Recollection rejection (RD produces retrieval of verbatim trace of similar target and correct rejection of RD)
F_D	Similarity judgment (RD produces retrieval of gist trace of similar target and acceptance of RD)
	True Memory Processes
E_T	Erroneous recollection rejection (target produces retrieval of verbatim trace for another target and incorrect rejection of target)
R_T	Identity judgment (target produces retrieval of verbatim trace and correct acceptance of target)
F_T	Similarity judgment (target produces retrieval of gist trace of similar target and acceptance of target)
	Response Bias Processes
B_V	URD produces false alarm in V condition
β_G	URD produces false alarm in G condition
β_{VG}	URD produces false alarm in VG condition

Brainerd, Stein, et al. (2008) applied the CR model to memory for emotional word lists in the DRM paradigm, becoming the first to measure retrieval processes involved in memory for emotional words. They used positive, negative, and neutral lists that had equivalent arousal levels and found differences in the CR parameters as well as differing levels of true and false memory. Negative word lists increased false memory compared to positive and neutral lists, because the familiarity parameter was higher and the recollection rejection parameter was lower. In other words, false memories for negative words were due to both increased gist and reduced verbatim retrieval for negative words.

Later, Bookbinder and Brainerd (2017) extended that sort of design to pictures. They applied a CR design to positive, negative, and neutral pictures with equivalent arousal levels in a DRM-like false memory design. Similar to Brainerd, Stein, et al.'s (2008) results with words, Bookbinder and Brainerd found that negative pictures increased false memory compared to positive and neutral pictures. In further parallel to Brainerd, Stein et al., this was because negative pictures increased familiarity and reduced recollection rejection.

Taken together these two sets of results suggest that negative emotional content elevates false memory through two mechanisms: increasing gist retrieval and reducing verbatim retrieval. However, this conclusion does not directly lead to predictions regarding mood because content and mood do not affect false memory in the same way. In the next section I discuss the theoretical reasons to expect different false memory and process effects as a result of mood compared to emotional content.

Theories of Emotion

In the preceding sections I have been referring to the distinction between emotional content and mood, and I have also briefly mentioned a second important emotional distinction:

valence versus arousal. Different theoretical approaches provide different terms for valence (e.g., approach/avoid, pleasant/unpleasant; Lang, Bradley, & Cuthbert, 1998; Tellegen, Watson, & Clark, 1999), but here I refer to the two dimensions as valence, ranging from negative to neutral to positive, and arousal, ranging from low to high. In Russell's (1980, 1991) circumplex model of emotion, valence and arousal are separate, continuous dimensions that are free to vary independently of each other. Any given emotion can be represented as a linear combination of valence and arousal. The independence of valence and arousal has been supported by neuroimaging research that has shown different patterns of brain activation for the processing of arousal (amygdala; Anderson, Christoff, Panitz, De Rosa, & Gabrieli, 2003) and valence (prefrontal and cingulate cortices; Colibazzi et al., 2010). This framework is also supported by data in which ratings of positivity and negativity are correlated, because in this conceptualization, positivity must decrease as negativity increases, and vice versa, because they are represented by the same scale (e.g., Russell & Carroll, 1999).

Others, however, have argued that valence and arousal are actually correlated. Instead of separate scales for valence and arousal, this framework posits two separate scales for positivity and negativity, with arousal represented on both scales (see Larsen, 2017). That is, as either positivity or negativity increase, arousal increases as well. This structure also allows for positivity and negativity to vary independently and has been supported by findings that valence and arousal are correlated in ratings of several kinds of stimuli (Bradley & Lang, 1999, 2007; Lang, Greenwald, Bradley, & Hamm, 1993) and that positivity and negativity can be uncorrelated (Cacioppo, Gardner, & Berntson, 1997).

In response to these two conflicting frameworks of emotion, Mattek, Wolford, and Whalen (2017) have recently developed a third framework that can potentially account for cases where positivity and negativity are correlated as well as when valence and arousal are correlated. In short, the relationships between valence and arousal and between positivity and negativity are not constant. Rather, valence and arousal my sometimes be correlated and sometimes be uncorrelated, and likewise for positivity and negativity. The determining factor, they argue, is how ambiguous valence is: when valence is unambiguous valence and arousal are correlated, and when valence is ambiguous positivity and negativity are inversely correlated. The model describing these relationships, taking into account valence ambiguity, accounts for much of the variance in valence ratings predicted by arousal. However, it leaves some questions unanswered, such as why the negativity-arousal correlation is stronger than the positivity-arousal correlation (see Brainerd & Bookbinder, under review). Therefore, there is still work to be done toward the development of a coherent framework of valence and arousal.

Manipulating Emotional Content and Mood

Both valence and arousal have been manipulated via both content and mood for the purpose of studying false memory. Although the findings are informative, I argue that a difficulty with the existing literature is that valence and arousal have not always been manipulated independently, in either content or mood research. In this section I briefly describe the way that emotional content is manipulated before discussing the way that moods are induced, which is of direct relevance to the present study.

Content. Emotional word lists are the most common method of altering the emotional content of target items. I discussed earlier the DRM paradigm as a method of measuring false memory; the neutral lists in the standard DRM design can be replaced with lists that vary in terms of valence, arousal, or both, in order to measure false (and true) memory for emotional items. Budson et al. (2006) were the first to develop said emotional DRM lists (e.g., tears, sad, tissue for the CD “cry”). However, due to limitations of Budson et al.’s lists such as the fact that they did not include positive lists and that arousal was not controlled, others have created lists in which arousal is controlled (Brainerd, Stein, et al., 2008) and factorially manipulated with valence (Brainerd, Holliday, Reyna, Yang, & Toglia, 2010). These improvements have allowed researchers to assess whether false memory can be affected by valence even when there is no influence of arousal and whether valence and arousal may interact to affect false memory. When valence and arousal were factorially manipulated, Brainerd et al. (2010) indeed demonstrated a valence-arousal interaction: Higher levels of arousal increased false memory, but only for negative lists.

Pictures have also been used as an emotional content manipulation by making use of picture databases that have been normed for valence and arousal. The IAPS (Lang, Bradley, & Cuthbert, 2005) and the GAPED (Dan-Glauser & Scherer, 2011) are two that contain a large number of color photographs with varying levels of valence and arousal. They also contain enough pictures with similar themes and meaning to enable their use in false memory paradigms, where pictures conveying the same meaning as targets are used as false memory items. Since these picture databases contain valence and arousal norms, valence and arousal can be controlled, although such control has not always been put into practice in prior work. The results of picture experiments are mixed because some have mirrored the pattern of word results (Bookbinder & Brainerd, 2017), whereas others have found that negative pictures reduce false memory (Choi, Kensinger, & Rajaram, 2013). After reviewing the results of mood induction studies, it will become clear that a potential explanation for Choi et al.’s result is that the pictures they used were powerful enough to affect subjects’ moods.

Mood. Various mood induction procedures (MIPs) have been implemented to alter subjects' emotional states before or after they are exposed to target items, which can be neutral or emotional themselves. An example of a mood induction protocol is the following: 1) pre-induction mood measurement, 2) mood induction, 3) post-induction mood measurement, 4) target presentation, 5) memory test. MIPs include videos (e.g., Gross & Levenson, 1995; Storbeck & Clore, 2011), music (Justin Storbeck, 2013; Justin Storbeck & Clore, 2005), a combined method that implements both music and a guided imagery task (Mayer, Allen, & Beauregard, 1995), or simply pictures (Storbeck, 2013). All of these MIPs can be implemented either before target presentation, in order to affect encoding processes, or after target presentation but before the memory test, in order to affect retrieval and consolidation processes. Before target presentation is the more typical procedure.

As a manipulation check, subjects' moods are measured before and after the mood induction, using a scale such as the Positive and Negative Affective Schedule (PANAS; Watson, Clark, & Tellegen, 1988), the Brief Mood Introspection Scale (BMIS; Mayer & Gaschke, 1988), or the Self-Assessment Manikin (SAM; Bradley & Lang, 1994). On the PANAS, subjects rate the degree to which they are experiencing a series of emotions (e.g., scared, joyful) at a particular time (for the purpose of a mood manipulation check, that time would be "right now," but other contexts may call for other time periods). On the BMIS subjects similarly rate how well several adjectives describe their present mood, and on the SAM they use a pictorial assessment to rate how happy/sad and how calm/excited they are currently feeling.

Although such mood inductions have been implemented in false memory designs, the results (reviewed below) have been called into question because of the conflation of valence and arousal. The successful disentanglement of two potential false memory factors is one of the primary motivations for the present research.

Emotion and Memory

In this section I discuss the existing theories that have been applied to explain emotional effects on true and false memory.

Emotional content. Turning to theories of emotional content first, one of the earliest accounts was Easterbrook's (1959) hypothesis, which proposed that emotional content focuses attention on central details and reduces attention to peripheral details, thereby increasing true memory for central details and reducing accuracy for peripheral details. The mechanism for the memory difference for different types of details was thought to be related to effects of arousal on attention. Hamann (2001) explained the adaptive nature of arousal-enhanced memory: Arousing events are likely to be related to survival, so improved memory for those events enables us to respond to them quickly or avoid risky events in the future. Although Hamann and Easterbrook not address valence, what is implicit in the survival-based account is that the arousing events are negative (i.e., risky). This explanation could not be applied to arousing but positive events. In terms of false memory, the implication is that when content is negative and arousing, false memories are more likely to occur for peripheral details of an event. This explanation is thought to account for a phenomenon known as the weapon focus effect (Loftus, 1979). Witnesses of crimes (highly emotional events), have highly accurate memory of the weapon (the central detail), but their memory for other details such as the appearance of the perpetrator (peripheral details) may not be as reliable.

Kensinger and colleagues (e.g., Kensinger, Garoff-Eaton, & Schacter, 2007) expanded this theory to events that were not necessarily crime-related, such as a picture of a snake in the forest, proposing that when events contain both emotional and neutral information, memory accuracy differs for the emotional versus neutral aspects of the event. As with the weapon focus effect, negative, arousing aspects of negative events are subject to improved memory compared to neutral aspects of negative events. False memory would be reduced for negative aspects of negative events compared to neutral aspects of negative events. Mather (2007) proposed an explanation of this effect on the basis of arousal: Emotional aspects of events will receive more processing, reducing false memory, because they are more arousing than the neutral content, which receives less processing. To sum up, initial accounts of emotional content effects centered on arousal to explain better memory for central, emotional information as compared to peripheral or neutral information, because arousal is an adaptive signifier of information that it is beneficial to remember.

Kensinger and colleagues have also proposed that valence plays a role as well, although their methodology does not necessarily directly lead to that conclusion. Kensinger and Corkin (2003) found that negative, arousing words elicit more "remember" responses than neutral words, although in their study it was possible that this effect was driven more by arousal than valence, since the negative words used were more arousing than the neutral words. Similarly, Kensinger, O'Brien, Swanberg, Garoff-Eaton, and Schacter (2007) found that subjects are better able to distinguish imagined information from presented information if it is negative relative to when it is positive or neutral. However, this effect is not purely related to valence, as negative items were more arousing than positive or neutral items. Thus, arousal in particular may improve memory for specific details of item's presentation, but the possibility of a valence contribution was not categorically ruled out. Also, it should be noted that this "remember" advantage for negative arousing stimuli is not indicative of higher accuracy. When Dehon, Larøi, and Van der Linden (2010) measured remember/know judgments for emotional DRM lists, they found more “remember” responses to CDs from negative lists compared to positive and neutral lists. Unlike some of Kensinger’s work, arousal was equivalent across valence, so Dehon et al. were able to conclude that it was negative valence per se that increased remember judgments. To sum up, negative or negative-arousing information may be associated with stronger feelings of remembering and better memory for details, but stronger feelings of remembering are not an indicator of accuracy in themselves.

The paradoxical negative emotion hypothesis (PNE; Porter, Taylor, & Ten Brinke, 2008) explains why negative information may elicit both improved accuracy and more distortion. As noted above, there may be an adaptive advantage to remembering negative information better in order to avoid future risks, but this increased memory may also involve the incorporation of more information and broader information, perhaps by consolidating information from several sources. This expanded memory for relevant negative information may result in higher risk of distortion because sources may be confused.

Others, however, have noted that positive information as well incurs a memory advantage over neutral, and sometimes negative information. Heuer and Reisberg (1990) proposed the idea of emotional enhancement of memory (EEM), which is the general idea that emotional content enhances memory (for both central and peripheral details) compared to neutral content. This effect may be related to both valence and arousal, as it is thought to be a result of both heightened physiological arousal and semantic relatedness, the latter of which is a feature of items of high emotional valence (Talmi & Moscovitch, 2004). Although the high semantic relatedness of emotional content increases true memory, it increases false memory as well, compared to neutral content. To truly disentangle EEM, Adelman and Estes (2013) carefully controlled valence and arousal levels in a large set of words and found that both positive and negative words were remembered better than neutral words regardless of arousal. Thus, there is additional evidence that valence contributes to memory independent of arousal.

Fuzzy-trace theory on emotion and memory. Clearly, the results are mixed when it comes to emotional content effects on memory, which may explain why a unified theory has yet to take hold. FTT provides an opponent process account for memory differences in emotional content, separately for valence and arousal, enabling it to account for many of the above effects. Turning to valence first, emotional content (whether positive or negative) should strengthen gist traces compared to neutral content because the semantic connections are stronger. Furthermore, this gist should be strengthened even more for negative valence compared to positive valence because the semantic connections for negative content are more salient. Thus, one would expect to find more false memories for emotional content compared to neutral content and to find more false memories for negative content compared to positive content. In terms of verbatim traces, they should be strengthened by positive valence compared to neutral valence and weakened for negative valence compared to neutral valence. Support for this claim comes from research on retrieval processes involved in recall of emotional words: Gomes, Brainerd, and Stein (2013) found better recall for positive words compared to negative words, which was because recall of positive words was more influenced by recollective retrieval, a process that involves reinstatement of contextual features of prior experience. On the other hand, recall of negative words was primarily accomplished through nonrecollective retrieval, which does not involve reinstatement of features of prior experience. Again, this result supports the prediction that negative content increases false memories, but contrary to the prediction about gist traces, suggests that positive content can also reduce false memories. Thus, positive content can increase or decrease false memory depending on whether it affects verbatim or gist traces, which may depend on methodological factors such as stimulus type.

FTT can thus explain predictions of both EEM and PNE accounts within a single theory. Furthermore, it makes predictions about arousal as well, whereas none of the earlier theories explicitly make predictions for both valence and arousal. Theories have tended to focus on either valence or arousal being the primary determinant of memory effects, but it might well be the interaction between valence and arousal that best explains memory effects. For arousal, the predictions focus on verbatim but not gist traces. A small increase in arousal will strengthen verbatim traces, but high levels of arousal will actually weaken verbatim traces because they blur attention. This prediction can be thought of as a variant of the Yerkes-Dodson law (Yerkes & Dodson, 1908), whereby performance increases with increasing arousal up to a certain point at which it plateaus and then declines. Thus, false memory would decline as arousal increases from low to moderate levels but would increase again when arousal becomes high enough. The various possible combinations of valence and arousal can thus lead to a variety of effects on false memory due to affecting verbatim and gist traces in different ways. In that connection, Gomes et al. (2013) found that the recall advantage for positive words compared to negative words held only when arousal was high; recall of negative words was dominated by nonrecollective retrieval when arousal was high but was dominated by recollective retrieval when arousal was low, suggesting that valence effects on word recall, at least, depend on the degree of arousal. The results also confirmed that true memory increases with moderately high levels of arousal, consistent with the above prediction concerning strengthened verbatim memory traces.

Associative accounts of emotional memory. Based on the assumption that valenced information is more densely connected that neutral information (Talmi, Luk, McGarry, &

Moscovitch, 2007), associative theories predict that valenced word lists are more associated, which leads to stronger activation of both studied words and any related words of the same valence. Thus, true and false memory should both be higher for emotional stimuli compared to neutral stimuli. In making separate predictions for positive and negative valence, some proponents of associative theories have argued that while both valences create stronger associations than neutral stimuli, negative stimuli are more distinct from each other and positive stimuli are more strongly associated with each other. Thus, negative valence should increase true memory but reduce false memory, compared to positive valence. In other words, memory *accuracy* would be better for negative information, due to the adaptive function of remembering threatening or survival-related information well (see Howe & Otgaar, 2013). Associative theories also predict valence-based differences in change in activation over time. Whereas activation of neutral related stimuli declines over time, activation of emotional stimuli, because it is more strongly associated, is more robust and can even increase over time (Howe, Candel, Otgaar, Malone, & Wimmer, 2010). Note that associative accounts do not make specific predictions about the role of arousal as it relates to activation of targets or CDs.

Theories of mood and memory. Bower's (1981) network theory of affect provided an early theoretical account that formed the basis for more recent theories. In this conception, mood provides a context for remembering so that when the mood experienced during encoding matches the mood experienced during retrieval, true memory will be enhanced. In other words, Bower proposed that memory will be enhanced when study and test moods are the same, a concept that is now known as mood-congruent memory. There is some recent evidence for such an effect on true memory as well as evidence of a mood-congruent memory effect on false memory (Knott & Thorley, 2013; Ruci, Tomes, & Zelenski, 2009). However, Bower did not make specific predictions about valence differences, nor did he directly address mood effects on false memory.

The affect-as-information hypothesis (AAI; Clore & Storbeck, 2006; Schwarz & Clore, 1983) was one of the first to address such questions. In terms of valence differences, this theory draws on the distinction between two types of processing defined by Hunt and Einstein (1981). Item processing, which focuses on distinguishing features of individual items, is employed under negative moods, whereas relational processing, which focuses on connections among items and their meanings, is employed under positive moods. Drawing on this distinction one would predict that positive moods promote false memory whereas negative moods protect against it, because false memories result from semantic connections but are inhibited by recollection of surface details unique to individual items. In this way, the item versus relational processing distinction is compatible with FTT's verbatim versus gist distinction, and considering the theories in tandem leads to clear predictions about false memory: Relational processing in positive moods encourages subjects to focus on the gist of target events, leading to false memory, but item-specific processing in negative moods shifts the focus to item's surface form, suppressing false memory by enhancing the ability to reject similar but not presented related distractors.

Within the affect-as-information framework, Clore and Storbeck (2006) further specified the conditions determining which processing style would be used. They argued that relational processing is the more dominant of the two styles, and that processing shifts to the item-specific style to signal a problem. Negative affect provides a signal that something is amiss, shifting processing to become more item-specific. For example, fear in response to a dangerous situation should stimulate item-specific processing of that situation because it will be advantageous to remember those details in the future.

Furthermore, Storbeck and Clore (2008) later integrated arousal into this framework, proposing that whereas negative affect may signal something dangerous or problematic, arousal signals importance or urgency. What is so powerful about the role of arousal in this framework is that the arousal need not be sourced from the object that is undergoing processing. Rather, arousal from an external source—e.g., physiological arousal—that is attributed to an emotional reaction to the processed object signals urgency of processing that object. Arousal can thus modulate judgment or memory via processing even if the item being judged or remembered is itself not arousing.

The classic example of this role of arousal is Dutton and Aron's (1974) bridge study, where men who crossed high bridges experienced high levels of arousal which they misattributed as physical attraction to a woman they met after crossing the bridge. More recently, Cahill, Gorski, and Le (2003) showed that increasing subjects' arousal via a cold pressor stress task increased memory for emotional but not neutral pictures. These results were interpreted as evidence that arousal signals the greater importance of negative stimuli, which resulted in enhanced encoding and subsequent improved memory.

Affect-as-information theory predicts that the mechanism for enhanced processing under high arousal is that higher levels of arousal signify high urgency, encouraging people to continue to use the processing style they are already using and elaborate more on their current mode of processing. In other words, high arousal intensifies reliance on item-specific processing in people experiencing negative moods, and high arousal intensifies reliance on relational processing in positive moods. Lower levels of arousal, however, should not impact processing style or could attenuate its effects. Evidence for this prediction on the negative side comes from findings that priming people with fearful faces in order to increase their arousal enhances item-specific processing as measured by contrast sensitivity in a discrimination task, but priming with neutral, non-arousing faces had no impact (Phelps, Ling, & Carrasco, 2006). Hurlemann et al. (2005) provided further evidence of arousal effects on both kinds of processing: When positive and negative "oddball" words were presented in a list of neutral words, negative words reduced memory for previous words whereas positive words enhanced memory for subsequent words, confirming that positive valence encourages relational processing and negative valence focuses processing on individual items at the expense of neutral items. When subjects' physiological arousal was increased, both of these effects became stronger, and when their arousal was lowered, both effects were reduced.

The affect-as-information approach does not explicitly make predictions about how arousal influences false memory, but if arousal is expected to amplify the effects of valence-directed processing style, then one would expect an interaction between valence and arousal. Specifically, false memory should be higher when subjects are in positive moods compared to negative moods and this difference should be larger when subjects' arousal is higher. In that connection, there is evidence of different processing styles associated with distinct negative moods (anger/high arousal and sadness/low arousal), suggesting that differences in arousal level (within a given valence) can affect memory via processing differences (Semmler & Brewer, 2002).

AAI and FTT make congruous predictions about valence, but are not quite as compatible regarding arousal. Whereas AAI predicts that arousal will modulate the effects of valence on memory, FTT predicts separate effects of valence and arousal, and thus arousal may affect memory without any impact of valence (although it can account for a modulating effect of arousal as well). Recall that arousal is expected to reduce false memory at low levels, but increase false memory at higher levels. Arousal may have such an impact across different levels of valence, or it may interact with valence; FTT allows for both options. Furthermore, the predicted valence-arousal interaction may not be the same for AAI and FTT. In AAI, higher arousal is expected to increase the false memory difference between positive and negative mood, whereas at low arousal the positive-negative difference would be smaller or even nonexistent. According to FTT, on the other hand, low levels of arousal should further reduce false memory for negative moods, but may reduce false memory in positive moods as well by strengthening verbatim traces. Higher levels of arousal would increase false memory in positive moods, but may also increase false memory in negative moods by reducing the ability to use verbatim traces. Thus, in FTT it is not clear whether valence-based false memory differences would be larger when arousal is low or high. One method of comparing the two theories, then, would be to compare valence differences in false memory when arousal is low versus high. If the valence gap increases, it would provide clear support for AAI, although it would not necessarily provide an argument against FTT. If valence differences are attenuated at higher levels of arousal, that is supportive of predictions of FTT and not AAI. Finally, if arousal is able to impact memory when valence does not have an effect, that would provide clear support for FTT.

Previous Finding and Gaps in Prior Research

A review of the existing research led to the conclusion that, in general, negative moods reduce false memory and the most straightforward explanation for this result is that mood induction primarily affects verbatim memory (Bookbinder & Brainerd, 2016). I review here some of the research that contributed to that conclusion that is most relevant to the present experiments. In an early experiment, Storbeck and Clore (2005) used music to induce positive and negative moods prior to a DRM task with neutral words. Consistent with predictions of FTT and affect-as-information theory, subjects in positive moods had higher levels of false recall and subjects in negative moods had lower levels of false recall compared to the control group. Storbeck and Clore (2011) expanded this result by comparing mood inductions before and after target presentation. They found that, again, negative mood reduced false memory when mood was induced before target presentation, but did not find false memory effects when mood was induced after. They concluded that mood affects encoding but not retrieval processes, consistent with both the processing style predictions of AAI and the memory trace storage predictions of FTT. Mood did not impact true memory in either study.

Storbeck (2013) provided further evidence that negative moods reduce false memory for neutral words, using both music and picture mood inductions. He found that negative moods reduced false recall and false recognition. When words were accompanied by pictures, however, negative mood no longer reduced false recognition compared to positive and neutral moods. Altogether these results support the prediction from the affect-as-information hypothesis that negative mood promotes item-specific processing, and supports the prediction from FTT that this processing produces stronger verbatim traces. The fact that the false memory result disappeared with pictures is consistent with these predictions because pictures strengthen verbatim traces, leaving less latitude for further strengthened by negative mood.

Emery, Hess, and Elliot (2012) and Knott, Threadgold, and Howe (2014) provided further corroborating evidence that negative moods reduce false recall, but arousal was not mentioned in either article, indicating that it was likely not controlled. Emery et al. used a video mood induction and found lower false recall for subjects in negative moods. Knott et al. did something slightly different and used their false recall task to prime responses to a compound remote associate task (CRAT) in order to confirm the hypothesis that false memories occur due to spreading activation. Subjects in the positive and neutral mood groups were better at solving CRAT problems when they had recalled the CD that was the solution to the problem. Subjects in the negative mood group, however, were not any better at solving the problems after the recall task. Knott et al. proposed that item-specific processing in negative moods leads to less spreading activation during priming. In terms of FTT, this result suggests that negative moods could enhance verbatim retrieval but also reduce gist retrieval, at least on certain tasks.

In all of the above studies, the false memory differences obtained could be alternatively explained by differences in arousal. In other words, the reduced false memory for the negative, relatively high arousal groups could be either due to item-specific processing or improved verbatim traces at moderate arousal levels. To pit valence against arousal in explaining mood effects on memory, Corson and Verrier (2007) factorially manipulated the two components of emotion. Following the same general procedure implemented by Storbeck and Clore, subjects' moods were induced prior to a DRM task with both recall and recognition tests. Five mood groups were created using music and guided imagery: positive/high arousal, positive/low arousal, negative/high arousal, negative/low arousal, and a control group. Corson and Verrier found effects of arousal but not valence, on both false recall and false recognition. Both forms of false memory were elevated for the high arousal groups compared to the low arousal groups, regardless of valence. Van Damme (2013) replicated these results, confirming that high arousal increases false recognition on immediate and delayed tests, but neither valence nor arousal impacted false memory on an immediate recall test. These results suggest that arousal may be a component of mood-false memory effects, and may indeed impact memory independent from valence, but considered alongside existing research, the relative contribution of valence and arousal is still an open question.

Most recently, Mirandola and Toffalini (2016) investigated mood effects on retrieval by inducing positive-high arousal, negative-high arousal, and neutral-low arousal moods immediately before a recognition test, using pictures from the IAPS. Instead of the DRM paradigm, they used a picture false memory task that encourages subjects to falsely remember non-presented aspects of scripted scenes. They measured false memory for causal errors (remembering an unpresented cause of a studied event) and gap-filling errors (remembering a script-consistent but unpresented detail). They found that negative and positive moods reduced causal errors compared to neutral moods, but mood valence did not affect gap-filling errors.

They then conducted a second analysis using subjects' mood valence and arousal ratings as opposed to the mood valence group they were assigned to as a factor. They found that arousal, but not valence, affected causal errors such that causal errors became less likely as arousal increased. Similarly, arousal but not valence affected gap-filling errors because these errors decreased as arousal increased. Finally, hits (true memory) were unaffected by valence but increased as arousal increased. Therefore, Mirandola and Toffalini found that when valence and arousal were disentangled using subjects' actual self-reported mood, arousal but not valence affected true and false memory. Higher levels of arousal resulted in greater memory accuracy as measured by both increased hits and reduced false alarms. Furthermore relative likelihood estimates of a model using self-reported valence and arousal compared to a model with mood group as a predictor revealed the self-report model to be a better predictor of memory performance. This new result calls into question prior research where subjective valence and arousal ratings were not factored into the false memory analysis. Although mood researchers always confirm that valence and arousal group differences are as intended, these groupings do not account for variability in subjects' moods within groups. In Mirandola and Toffalini's population, at least, this variability was better at predicting mood effects than preassigned mood groups.

The experiments reviewed in this section provide evidence for two main results: a) that negative mood reduces false memory when arousal is not controlled (and thus arousal is likely higher for the emotional groups compared to the neutral group and may also higher for the negative group than the positive group) and b) that high arousal increases false memory and the valence effect disappears when valence and arousal are factorially manipulated. Although Mirandola and Toffalini found the opposite effect of arousal, the fact that their MIP occurred before retrieval makes this result difficult to compare to the others. The simplest explanation for the main pair of results is that mood affects verbatim memory, because it is strengthened by negative valence and weakened by high arousal. It is not clear from these results, however, whether valence effects on false memory hold when arousal is held constant across all groups. If they do, then arousal is not a necessary component of mood-false memory effects.

Furthermore, no existing research has addressed this hypothesis—that mood affects verbatim memory in order to influence false memory—directly by measuring retrieval processes. One can postulate that the mechanism is likely related to verbatim retrieval based on existing research, but in order to confirm this prediction it is necessary to directly measure the contributions of verbatim and gist retrieval to false memory as a function of mood. By using the conjoint recognition model, the present research is the first to do so as well as identify which verbatim and gist processes can explain differences in false memory rates.

Goals of the Present Research

Following the evidence from prior research that both valence and arousal may impact false memory depending on how the two factors are manipulated, the present research was designed to answer four questions: 1) Can purely valence-based mood effects be obtained when arousal is moderate, and which processes contribute to mood differences in true and false memory? 2) Can mood effects be strengthened by encouraging verbatim processing of study lists, indicating that mood creates a primarily verbatim effect? 3) Do mood valence effects change when arousal is increased? and 4) Can arousal affect false memory independent form valence, and what processes contribute to that effect?

The first question was answered in Experiment 1 by implementing a mood induction that kept arousal at a low level, equivalent across positive, negative, and neutral moods. The short answer is that mood valence was effective at influencing false memory, even when arousal was low, because negative moods increased verbatim retrieval. The second question was answered in Experiment 2 by adding a manipulation to differentiate verbatim and gist processing. The addition of this manipulation resulted in different processes causing the mood effect depending on whether verbatim or gist retrieval was favored. The third question was answered in Experiment 3 through the use of a different mood induction creating higher arousal. Elevating arousal changed the pattern of mood effects, increasing false memory in negative moods where false memory had been previously low. Finally, the fourth question was answered in Experiment 4 by measuring false memory as a function of arousal, independent from valence. Arousal was indeed able to impact false memory even when valence was controlled and this effect was driven by gist processing.

Structure of the Experiments

Chapter 2 details the first step in the research: the development of the mood induction procedures used in the three main experiments. It includes three pilot tests based on procedures used in prior research. Pilot 1 was a music and guided imagery approach which failed to produce reliably different moods in subjects. Pilot 2 used videos that produced reliably different levels of valence at equivalent low arousal levels. Finally, Pilot 3 followed a similar video procedure to Pilot 2 but in this case arousal was manipulated as well. Therefore, the MIPs developed in Pilot 2 and 3 were used for the main experiments.

The four main experiments, detailed in Chapters 3-6, all followed a similar procedure. Subjects' moods were measured, they viewed the videos, and their moods were measured again. The rest of the procedure, in all four experiments, followed a typical DRM design combined with conjoint recognition. That is, subjects listened to a series of lists and completed a recognition test with specific instructions after all lists had been presented. Additionally, this design included a second memory test after a delay period of two days (see Bookbinder & Brainerd, 2017), which enabled the inclusion of some additional analyses to provide further insight into what was happening at a processing level. All experiments followed this general procedure and the data from all experiments were analyzed in the same way. Subjects were excluded from the analysis if their moods were inconsistent with the manipulation and the data were bias-corrected using responses to URD items. Analyses of variance (ANOVAs) were conducted on the acceptance rates of targets and CDs, and the conjoint recognition model was fit to the acceptance data and its parameters were estimated. A complete picture of the data was provided by interpreting the results in terms of both ANOVAs and parameter estimates, and connections were made between findings from the two analysis methods.

Hypotheses, Predictions, and Theoretical Framework

Experiment 1 was designed to test the hypothesis that arousal is not a necessary component of emotion-memory effects. I predicted that true memory would not be affected by valence based on the lack of TM effects in the literature (Corson & Verrier, 2007; Storbeck & Clore, 2005). However, I predicted that valence could affect false memory with arousal held at a moderate level. Previous research has persisted in the claim that mood valence should affect false memory based on the item-specific versus relational processing distinction. Although Corson and Verrier (2007) and Van Damme (2013) provided some evidence that valence does not affect false memory when arousal is controlled at high or low levels, they did not address whether valence effects could occur at moderate arousal levels. Furthermore, very high arousal is expected to increase false memory by blurring attention, and could potentially create a ceiling effect such that valence differences would be undetectable. On the other hand, when arousal is very low, differences in valence are likely to be smaller. Therefore, if pure valence effects are to be detected, arousal must be held at a moderate level. If negative mood encourages item-specific processing or affects verbatim or gist retrieval, it should be the case that false memory will be reduced for negative, moderately aroused moods compared to positive moods. I predicted that false memory rates would thus be lower for subjects in negative moods, and more specifically that this difference would result from increased verbatim processes (e.g., recollection rejection) rather than reduced gist processes (e.g., familiarity).

Experiment 2 was designed to test the hypothesis that mood-memory effects are dependent upon strong verbatim memory traces. Shorter DRM lists were included to determine whether mood-false memory effects would be stronger when verbatim traces are more accessible, compared to longer lists where gist traces would be favored. I predicted that negative valence would reduce false memory for short lists, which would indicate that a) pure valence can affect false memory under certain conditions and that b) it is a verbatim effect. Not finding any valence differences would support the claim that valence is not sufficient to produce mood effects on false memory even when verbatim traces are strong, at least at moderate arousal levels.

Experiment 3 addressed the hypothesis that increasing arousal would attenuate the difference in false memory between positive and negative moods. This hypothesis is supported by prior research showing that emotional effects on false memory are the result of an interaction between valence and arousal, at least for emotional content (Gomes et al., 2013), or are wholly dependent on arousal (Corson & Verrier, 2007). If pure valence effects were detected in the first two experiments, increased arousal could interact with valence and provide a test of the competing hypotheses born out of AAI and FTT: Higher arousal increasing the false memory difference between positive and negative valence would support AAI, whereas higher arousal reducing that difference would be more consistent with FTT.

Finally, Experiment 4 addressed the hypothesis that high levels of arousal can increase false memory regardless of valence, and that this effect is driven by reduced verbatim retrieval at higher levels of arousal. In other words, this study was designed to determine whether arousal is sufficient to produce false memory effects. This prediction is based on FTT's account of arousal, wherein it enhances verbatim retrieval at low levels but reduces verbatim retrieval at high levels, and this verbatim effect does not rely on processing styles associated with different valences. In this final study, arousal was measured based on subjects' self-report, in accordance with Mirandola and Toffalini's (2016) finding that these ratings are a better predictor of false memory. If arousal is found to increase false memory here, it would provide corroborating evidence for the claim that higher arousal increases false memory and would additionally provide information on the retrieval processes controlling that effect. Furthermore, a pure effect of arousal is supportive of FTT, but not, AAI where the influence if arousal is only to modulate valence effects.

Other manipulations. Beyond the major goals of the experiments outlined above, I included several other manipulations relevant to emotionally-laden real-life situations. In all four studies, I used a form of memory test that provided information on two additional factors that affect true and false memory. The general form of this type of recognition test is as follows. On the test that immediately follows the study phase, subjects respond to only half of the memory test items. On the later memory test, subjects respond to all memory test items; that is, over the course of the whole procedure they respond to half of those items twice and they respond to the other half only once. In this way I was able to compute measures of two useful factors: retention interval, which provides information on forgetting over time, and repeated testing. Retention interval can be measured by comparing the immediate test items to half of the delayed test items—the ones that were not previously tested on the immediate test. Any difference in memory for immediate test versus delayed test items is due to any forgetting that occurs before the items are tested. Repeated testing is measured by comparing the two halves of items on the delayed test. One half was already tested once before and the other half was tested for the first time. Any difference in memory for those two groups of items is due to prior testing effects.

The inclusion of these two factors was relevant for several reasons. First, both of these factors provide clear predictions about verbatim and gist memory and so any general forgetting or repeated testing effects can provide confirmation of those theoretical predictions. More specifically, repeated testing should increase both true and false memory if it increases both verbatim and gist memory, but forgetting should cause a reduction in both true and false memory, although that reduction should be smaller for false memory, if it reduces verbatim memory more than gist memory (McDermott, 2006; Roediger & McDermott, 1995; Thapar & McDermott, 2001).

Second, if forgetting and repeated testing are clearly linked to increases or decreases in gist or verbatim retrieval, any differences in their effects on subjects in different moods can provide more information about whether verbatim or gist retrieval is emphasized by a particular mood. For instance, if it is true that forgetting causes a reduction in verbatim traces but not gist traces, and forgetting has a larger impact on memory in negative moods than positive moods, there is support for the hypothesis that negative moods encourage verbatim processing.

Finally, repeated testing and forgetting are relevant to situations when witnesses of crimes or people who have experienced abuse are asked the same questions more than once or are not asked until some time has passed since the original event. Knowledge about how repeated testing and forgetting affect false memory—especially when subjects are in negative moods—would be important to consider in the context of police interviewing, eyewitness testimony, therapy, medical reporting, etc.

The primary goals of my dissertation were to determine the impact of mood valence and arousal on true and false memory, to identify the retrieval processes associated with mood-memory effects, and to identify conditions under which mood effects on false memory are most likely to occur.

CHAPTER 2

PILOT TESTS: DEVELOPMENT OF MOOD INDUCTION PROCEDURE

In order to develop a mood induction procedure that could be effectively implemented online, I pilot tested two previously-used MIPs adapted to be presented as part of an online survey. The first method was Mayer, Allen, and Beauregard's (1995) music and guided imagery induction. The second was the presentation of videos, as done by Storbeck and Clore (2011), for example. For the music and guided imagery and the first video inductions, arousal was intended to be held constant across three mood groups differing in valence. For the second video induction, arousal was intended to be higher for the positive and negative mood groups compared to the neutral group, but equivalent across the positive and negative groups.

Pilot 1: Music and Guided Imagery (Low Arousal)

Mood inductions that have ostensibly been designed to manipulate subjects' valence have often tended to alter subjects' arousal as well, making it impossible to determine whether valence or arousal cause any resulting effects. Researchers have noted the difficulty of maintaining equivalent arousal levels across positive, negative, *and* neutral moods groups, because neutral moods cannot be very arousing, and as positivity and negativity become stronger arousal is apt to increase as well (e.g., Bradley & Lang, 1999). However, the challenging task of creating a mood induction that controls arousal is necessary in order to determine whether mood effects on memory can be obtained as a function of valence alone. I began with the combined music and guided imagery technique because of its past efficacy (Corson & Verrier, 2007) as well as the fact that the wide range of normed musical stimuli allowed me to select pieces with high degrees of positivity and negativity but low arousal.

Method

Subjects. 29 undergraduate students participated for course credit (M_{age} = 19.34, *SD* = 1.79, 14 women). 9 subjects were in the positive mood condition, 10 were in the negative mood condition, and 10 were in the neutral mood condition.

Materials. The SAM (Bradley & Lang, 1994) and the PANAS (Watson, Clark, & Tellegen) were used to measure subjects' moods before they completed the mood induction task and after. They rated their current levels of valence and arousal on two 9-point scales ranging from calm to excited (arousal) and unhappy to happy (valence). They also rated how much they were currently experiencing each of 16 emotions on a 4-point scale from "definitely do not feel" to "definitely feel."

The materials for the mood induction task were comprised of 1) musical pieces and 2) one-sentence descriptions of positive, negative, or neutral situations. Three musical pieces were used for this task: one was positive, one was negative, and one was neutral. Selection of these pieces was based on norms found in Eerola and Vuoskoski's (2010) database. They were each about three minutes long. 24 sentences were used: 8 were positive (e.g., "you have worked hard on a paper for the past month and you just found out you got an A on it"), 8 were negative (e.g., "your friend cancels your plans for tonight at the last minute"), and 8 were neutral (e.g., "the capital is usually one of the biggest cities in the country").

Table 2.1

Descriptive Statistics for Pilot Tests

		Pilot 1		Pilot 2		Pilot 3	
		M	*SD*	*M*	*SD*	*M*	*SD*
ArousalPre	Positive	4.67	1.41	4.12	1.62	4.65	1.57
	Negative	3.90	1.52	4.20	1.23	4.23	1.60
	Neutral	4.00	1.41	4.14	1.75	4.43	1.47
ValencePre	Positive	6.22	1.09	4.94	1.95	6.10	1.17
	Negative	6.60	1.17	5.70	1.42	6.41	1.14
	Neutral	6.20	1.62	5.71	1.38	5.81	1.08
ArousalPost	Positive	5.67	1.73	3.94	1.52	5.60	1.70
	Negative	5.00	1.41	4.20	0.79	5.09	1.38
	Neutral	4.90	1.60	4.07	1.69	4.14	1.85
ValencePost	Positive	6.44	1.01	6.00	2.00	6.45	1.32
	Negative	5.80	1.32	4.30	1.34	3.68	1.17
	Neutral	5.80	1.87	5.35	1.66	6.10	1.00
ArousalDiff	Positive	1.00	1.41	-0.18	1.70	0.95	1.15
	Negative	1.10	1.20	0.00	1.63	0.86	2.21
	Neutral	0.90	1.20	-0.07	1.59	-0.29	1.90
ValenceDiff	Positive	0.22	0.67	0.41	2.43	0.35	0.88
	Negative	-0.80	1.75	-1.40	1.51	-2.73	1.64
	Neutral	-0.40	1.90	0.29	1.68	0.29	0.85

Procedure. Subjects completed the pilot test online through the survey website Qualtrics (Qualtrics, Provo, UT). After completing the first mood measurement, subjects were instructed that they should try to enter the given mood state (positive, negative, or neutral) and to press a button when they were ready. After the button press, the music started to play and they were again instructed to think about how pleasant or unpleasant it was (no instruction was given in the neutral condition). After listening to the music for 1 minute, subjects received more instructions.

> "Now, the music will continue as you read several sentences. You will have some time to focus on each sentence before proceeding to the next. Please use this time to imagine what you would feel like in the situation described by the sentence, and allow yourself to feel the [positive or negative] feelings associated with it. Feel free to re-read the sentence as many times as you are able. The survey will automatically proceed to the next sentence once the allotted time is up. Click to continue when you are ready."

After pressing the button, the music resumed and the first sentence was presented. It stayed on the screen for 30 seconds before proceeding to the next sentence. The music continued while all 8 sentences were displayed (about 4 minutes; when the song ended it restarted immediately). After the last sentence was displayed the music stopped and subjects completed the second mood measurement. The entire task took about 10-15 minutes to complete.

Results

Difference scores were computed for valence and arousal to determine the change in mood as a result of the induction. One-way ANOVAs were conducted on pre- and post-induction valence and arousal as well as on the difference scores. Pre-induction mood measurement results revealed no reliable differences in valence or arousal for the three groups; $F(2,28) = .28$ for valence and $F(2,28) = .77$ for arousal, both *ns*. However, post-induction mood scores did not differ by valence, $F(2,28) = .61$, or arousal, $F(2,28) = .65$, both *ns*; difference scores did not differ by valence, $F(2,28) = 1.02$, or arousal, $F(2,28) = .06$, either, both *ns*. These results show that despite its effectiveness in prior research, the music and guided imagery mood induction was not effective in the present sample (see Table 2.1), perhaps due to its implementation online rather than in person. However, others have previously demonstrated the efficacy of certain types of MIPs online, so the lack of results was more likely due to the particular MIP used (Verheyen & Göritz, 2009).

Pilot 2: Video Mood Induction (Low Arousal)

Following Storbeck and Clore's (2011) successful use of a video MIP in a false memory design, this pilot test was designed to test the efficacy of video mood induction online.

Method

Participants. 41 undergraduate students participated in exchange for course credit ($M_{age} = 19.87$, $SD = 1.54$, 32 women). 17 subjects were in the positive mood condition, 10 were in the negative mood condition, and 14 were in the neutral mood condition.

Materials. Three short videos were used for the pilot test. The selection of videos was based on Gabert-Quillen, Bartolini, Abravanel, and Sanislow's (2015) norms for emotional film clips. In the neutral condition, the video depicted several scenes from nature and human exploration with peaceful background music. The selection of this video was based on Gabert-Quillen et al.'s use of a video taken from the introduction to the television series *Planet Earth* as a neutral stimulus. The negative video was a scene from *The Shawshank Redemption*, based on Gabert-Quillen et al.'s norms for this video ($M_{pleasantness} = 2.22$; $M_{arousal} = 2.62$) on a 5-point scale. Finally, the positive video depicted babies and puppies. Each video was approximately three minutes long.

Procedure. Subjects completed the pre-induction SAM and BMIS measures, viewed the videos, and completed the post-induction SAM and BMIS measures. They were given the following instructions prior to watching the videos:

> "In the next part of the survey, you will watch a short video. Please pay close attention to the images and sounds and allow yourself to feel any positive/negative emotions induced by the video. (Previous sentence omitted in neutral condition). Try to imagine yourself in the context of the video and feel any emotions that arise."

Results

Difference scores were computed for valence and arousal to determine the change in mood as a result of the induction. One-way ANOVAs were conducted on pre- and post-induction valence and arousal as well as on the difference scores. Pre-induction mood measurement results revealed no reliable differences in valence or arousal for the three groups; $F(2,38) = 1.07$ for valence and $F(2,38) = .01$ for arousal, both *ns.* However, post-induction mood scores did marginally differ by valence, $F(2,38) = 2.77$, $p = .08$, but not arousal, $F(2,38) = .10$, *ns*, as predicted.

Furthermore, difference scores were marginally affected by valence, $F(2,38) = 2.94$, $p = .07$, but not arousal, $F(2,38) = .04$, *ns*. Planned comparisons revealed significant differences between valence difference scores for the negative group compared to the neutral group, $p < .01$, for the positive group compared to the negative group, $p < .01$, and inspection of the means revealed a trend in the predicted direction for the comparison between the positive and neutral groups (see Table 2.1 for descriptive statistics). Because the pattern of means was appropriate and the effects were marginal, I concluded that differences in mood would become reliable when a larger sample was used in the primary experiments.

Pilot 3: Video Mood Induction (High Arousal)

Following Pilot 2 resulting in a successful mood induction in which arousal levels were held equivalent, I then developed a second set of videos in which the arousal levels of the emotional conditions would be higher—but not extremely high—in the effort to create stronger memory effects.

Method

Participants. 63 undergraduates participated in exchange for course credit (M_{age} = 20.10, *SD* = 1.35, 46 women). 22 subjects were in the negative-high arousal group, 20 were in the positive-high arousal group, and 21 were in the neutral group.

Materials. Two new high-arousal positive and negative film clips were selected based on Schaefer, Nils, Sanchez, and Philippot's (2010) norms for emotional film clips. The neutral film clip was the same as the one used in Pilot 2. The negative video was a scene from *American History X*, based on Shaefer et al.'s norms for this video ($M_{negativeaffect}$ = 2.73[1]; $M_{arousal}$ = 5.84) Finally, the positive video was a scene from *The Dead Poet's Society* based on Schaefer et al.'s norms ($M_{positiveaffect}$ = 2.82; $M_{arousal}$ = 5.66). These values represent moderately high levels of arousal and high levels of negative and positive affective tone, respectively.

Procedure. The procedure was the same as the procedure for Pilot 2.

Results

Difference scores were computed for valence and arousal to determine the change in mood as a result of the induction. One-way ANOVAs were conducted on pre- and post-induction valence and arousal as well as on the difference scores. Pre-induction mood

[1] Negative and positive affect were measured on separate 1-3 scales, so these values represent fairly high levels of valence. Arousal was measured on a scale from 1-7.

measurement results revealed no reliable differences in valence or arousal for the three groups; $F(2,60) = 1.27$ for valence and $F(2,60) = .39$ for arousal, both *ns*. However, post-induction mood scores did differ by both valence, $F(2,60) = 35.71$, $p < .01$, and arousal, $F(2,60) = 4.15$, $p < .05$.

Furthermore, difference scores were affected by valence, $F(2,60) = 46.74$, $p < .01$ and marginally differed for arousal as well, $F(2,60) = 2.99$, $p = .06$. Planned comparisons revealed significant differences between valence difference scores for the positive group compared to the negative group, $p < .01$, and for the negative group compared to the neutral group, $p < .01$, and inspection of the means revealed a trend in the predicted direction for the comparison between the positive group and the neutral group (see Table 2.1 for descriptive statistics). Planned comparisons also revealed significant differences between the arousal levels of the two emotional groups and the neutral group, $p < .05$ for both comparisons, but not between the two emotional groups, as expected.

CHAPTER 3

EXPERIMENT 1: EFFECT OF MOOD VALENCE ON TRUE AND FALSE MEMORY FOR NEUTRAL WORDS

Findings in the mood-false memory literature have been interpreted as evidence that negative moods reduce false memory compared to positive moods. However, issues with control of arousal suggest that we should question the validity of that claim. It is possible that instead, arousal differences between valenced and neutral moods can account for the variability in false memory. To determine whether valence truly impacts false memory, independent of arousal, the present study implemented a MIP with low arousal across all three valence groups. Beyond the typical measures of true and false memory, this study also identified retrieval processes associated with mood valence and its effect on memory as a function of forgetting and repeated testing.

Method

Participants

117 subjects participated in exchange for course credit. 91 subjects were women and all were between the ages of 18 and 25 (M_{age} = 20.52, *SD* = 1.29). 35 subjects were in the negative condition, 51 were in the neutral condition, and 31 were in the positive condition. Conjoint recognition instruction was manipulated within-subjects.

Materials

The mood induction materials were identical to those used in Pilot 2. They included the pre- and post-induction mood measurements and the three videos designed to induce positive, negative, and neutral moods with equivalent levels of arousal.

The memory task materials consisted of 234 words. Based on Stadler, Roediger, and Mcdermott's (1999) norms for false memory of 36 standard DRM lists, the 18 lists with the highest levels of recognition were selected for use in the present experiments. The one exception was that the “anger” list was not used and the next best list was used in its place, to avoid presenting inherently emotional words. The 11 words with the highest combined BAS and FAS values to the critical word were identified. The top ten of these words served as targets and the 11th served as a related distractor for each list. The decision to use 10-item word lists was based on Robinson and Roediger’s (1997) norms for false recognition as a function of number of target items. 18 unrelated distractors were selected from the 18 unused lists. Thus, the materials were comprised of 18 CDs, 18 RDs, 18 URDs, and 180 targets. See Table 3.1 for lists of targets. Audio files were created for presentation of the 18 10-word lists. The words were read aloud in order of descending associative strength at a two-second rate.

Table 3.2 displays how the conjoint recognition test questions were assigned to the items on the immediate and delayed test. As shown, on the immediate test subjects responded to 9 CDs, 9 RDs, and 9 URDs, with three of each item type being assigned to each of the three questions. They also responded to 18 targets, with 6 assigned to each of the three questions. There were 72 items on the immediate test. On the delayed test, all of the values were doubled, such that subjects responded to 18 CDs, 18 RDs, 18 URDs, and 36 targets, with 6 of each question for CDs, RDs, and URDs, and 12 of each question for targets. In this way, all of the immediate test items were tested again on the delayed test, along with an equal number of items that not been tested before. Items that were being tested for a second time were paired with the same CR question that they had been paired with originally. All 18 of the CDs, RDs, and URDs were thus tested on the delayed test, along with 36 of the 180 targets, comprising 144 items on the delayed test.

Table 3.1

Experiment 1 Items

CD	*doctor*	*needle*	*sleep*	*smoke*	*foot*	*sweet*	*trash*	*mountain*	*cup*
T1	nurse	thread	bed	cigarette	toe	sour	garbage	hill	saucer
T2	physician	syringe	nap	cigar	shoe	candy	dump	climber	mug
T3	stethoscope	pin	rest	pipe	inch	sugar	can	climb	glass
T4	surgeon	haystack	tired	fire	ankle	bitter	litter	peak	Measuring
T5	patient	sewing	snooze	chimney	hand	honey	junk	valley	coffee
T6	dentist	injection	drowsy	puff	leg	tart	waste	summit	tea
T7	medicine	thimble	slumber	flame	sandal	chocolate	landfill	range	soup
T8	lawyer	prick	dream	billow	sock	taste	bag	ski	drink
T9	sick	knitting	snore	blaze	boot	tooth	rubbish	molehill	straw
T10	health	eye	wake	lung	kick	heart	debris	steep	Sip
RD	cure	sharp	yawn	tobacco	knee	nice	pile	top	coaster
UD	shirt	lion	thief	music	army	rubber	king	car	fruit

CD	*river*	*city*	*chair*	*cold*	*window*	*smell*	*soft*	*rough*	*slow*
T1	stream	town	table	hot	pane	aroma	hard	smooth	fast
T2	creek	suburb	seat	chilly	sill	scent	pillow	sandpaper	snail
T3	flow	state	couch	shiver	shutter	stench	gentle	tough	turtle
T4	lake	county	sit	arctic	glass	sniff	smooth	bumpy	quick
T5	brook	country	stool	frigid	door	fragrance	cotton	uneven	molasses
T6	water	street	desk	ice	curtain	nose	skin	coarse	speed
T7	boat	village	sofa	freeze	ledge	hear	feather	gravel	delay
T8	swim	metropolis	bench	warm	shade	see	plush	ground	traffic
T9	fish	capital	cushion	frost	view	nostril	loud	jagged	lethargic
T10	tide	urban	rocking	heat	open	breathe	tender	sand	sluggish
RD	bridge	big	recliner	winter	house	stink	touch	rugged	stop
UD	black	girl	pen	flag	angry	spider	high	man	bread

Table 3.2

Experiment 1 Test Items with CR Statement

Type	*Word*	*Times Tested*	*CR*	*Type*	*Word*	*Times Tested*	*CR*
CD	foot	1	G	CD	chair	2	G
CD	mountain	1	G	CD	rough	2	G
CD	cup	1	G	CD	slow	2	G
CD	needle	1	V	CD	river	2	V
CD	sweet	1	V	CD	cold	2	V
CD	trash	1	V	CD	smell	2	V
CD	doctor	1	VG	CD	city	2	VG
CD	sleep	1	VG	CD	window	2	VG
CD	smoke	1	VG	CD	soft	2	VG
RD	cure	1	G	RD	coaster	2	G
RD	sharp	1	G	RD	bridge	2	G
RD	nice	1	G	RD	rugged	2	G
RD	knee	1	V	RD	winter	2	V
RD	top	1	V	RD	stink	2	V
RD	big	1	V	RD	stop	2	V
RD	yawn	1	VG	RD	recliner	2	VG
RD	tobacco	1	VG	RD	house	2	VG
RD	pile	1	VG	RD	touch	2	VG
T1	eye	1	G	T1	urban	2	G
T1	kick	1	G	T1	rocking	2	G
T1	sip	1	G	T1	sand	2	G
T1	health	1	V	T1	open	2	V
T1	heart	1	V	T1	tender	2	V
T1	tide	1	V	T1	sluggish	2	V
T1	wake	1	VG	T1	steep	2	VG
T1	lung	1	VG	T1	heat	2	VG
T1	debris	1	VG	T1	breathe	2	VG
T2	physician	1	G	T2	chilly	2	G
T2	candy	1	G	T2	scent	2	G
T2	suburb	1	G	T2	sandpaper	2	G
T2	nap	1	V	T2	mug	2	V
T2	shoe	1	V	T2	seat	2	V
T2	dump	1	V	T2	snail	2	V
T2	syringe	1	VG	T2	creek	2	VG
T2	cigar	1	VG	T2	sill	2	VG
T2	climber	1	VG	T2	pillow	2	VG
UD	shirt	1	G	UD	pen	2	G

UD	music	1	G	UD	angry	2	G
UD	fruit	1	G	UD	man	2	G
UD	rubber	1	V	UD	spider	2	V
UD	king	1	V	UD	high	2	V
UD	car	1	V	UD	bread	2	V
UD	lion	1	VG	UD	black	2	VG
UD	thief	1	VG	UD	girl	2	VG
UD	army	1	VG	UD	flag	2	VG

Procedure

Subjects participated in the experiment online through the survey platform Qualtrics in exchange for course credit. They completed the experiment in two separate sessions separated by approximately 48 hours. The first session included the mood induction, the study phase, and the immediate conjoint recognition test. The second session was the delayed conjoint recognition test.

To begin the first session, subjects underwent the mood induction identical to the second pilot test: They completed the pre-test mood measurement, viewed one of the three videos, then completed the post-test mood measurement. Then, they proceeded to the study phase, in which they listened to all 18 auditory word lists. The order of the lists was randomized, although the order of items within each list was held constant in order of descending BAS. Subjects were instructed to simply listen to the lists. After each list was presented, subjects were given 5 seconds to work on an arithmetic problem before the next list was presented. After the final list was presented, subjects were given one minute to work on as many arithmetic problems as they could complete.

After one minute they proceeded to the first (immediate) conjoint recognition test. The instructions explained that they would be presented with a series of words that may or may not have been presented previously. Each word would be accompanied by one of three statements:

1) “I heard this word during the study phase;” 2) “This word is new but similar to a word from the study phase;” or 3) “I heard either this word or a similar word during the study phase.” Subjects were instructed to respond to each word by clicking a button to indicate that the provided statement was correct or incorrect. They were then provided with several examples explaining how to respond to the statements before proceeding to the memory test. Subjects were able to spend as much time as they needed responding to each item, and they were required to provide an answer (“yes, correct” or “no, incorrect”) to each item.

Data Analysis

Responses on the recognition test were bias-corrected using the two-high threshold method (Snodgrass & Corwin, 1988). ANOVAs were conducted on the bias-corrected target and CD acceptance rates, separately for the immediate and delayed tests. Analyses were first restricted to the V condition as a measure of standard true and false memory, and were then extended to the full data set (i.e., all three instructional conditions). Additional ANOVAs were conducted on the URD acceptance rates to determine whether false alarms were influenced by mood.

Finally, estimates of the conjoint recognition model were computed after confirming that the model provided an adequate representation of the data. The model was fit using the responses to the three instructional conditions across all factors in the design, and parameter significance tests were computed to assess theoretically-motivated parameter differences in different conditions. This method was followed for all four experiments.

Results

Mood manipulation check

Participants whose mood was not manipulated as intended were excluded from the analyses, resulting in 117 subjects (35 in the negative mood group, 51 in the neutral mood group, and 31 in the positive mood group). Participants were excluded if the change in their valence rating from the pre-test to the post-test was more than 2 points in the direction opposite to the intended direction (e.g., participants in the positive group whose moods became substantially more negative), and participants were excluded if the change in their arousal rating from the pre-test to the post-test was more than two points (i.e., if there was a substantial increase in arousal). Analysis of the resulting 117 subjects revealed reliable differences in post-test valence for the three groups, $F(2,116) = 37.61$, $p < .05$, with valence being the highest for the positive group ($M = 7.28$, $SD = .76$), followed by the neutral group ($M = 5.52$, $SD = .59$), followed by the negative group ($M = 3.49$, $SD = .54$). As expected, there were no reliable differences in arousal among the three groups, $F(2,116) = .86$, *ns*, ($M_{\text{positive}} = 4.59$, $M_{\text{neutral}} = 4.19$, $M_{\text{negative}} = 4.31$).

Bias correction

The effects of CR question and mood on unrelated distractor acceptance were identified in order to determine the proper method of bias correction. Two parallel 3 (question: V, G, VG) x 3 (mood: positive, negative, neutral) ANOVAs were performed on the acceptance rates for unrelated distractors, separately for the immediate and the delayed tests. There were reliable main effects of question on both tests and a reliable effect of mood on the immediate test but no interactions (see Table 3.3 for descriptive statistics). On the immediate test, unrelated distractor acceptance was higher in the G condition than the V and VG conditions, $F(2,228) = 9.67$, $p < .05$.

Table 3.3

URD Acceptance Rates

		Experiment 1		Experiment 2		Experiment 3	
		M	*SD*	*M*	*SD*	*M*	*SD*
				Immediate			
	Negative	.21	.27	.26	.33	.38	.24
V	Neutral	.24	.28	.39	.36	.31	.30
	Positive	.26	.27	.29	.32	.24	.30
	Total	.24	.27	.32	.34	.30	.31
	Negative	.31	.30	.45	.29	.41	.33
G	Neutral	.37	.34	.50	.30	.35	.29
	Positive	.47	.31	.39	.31	.41	.29
	Total	.38	.33	.45	.30	.39	.30
	Negative	.16	.27	.22	.28	.22	.29
VG	Neutral	.27	.31	.21	.30	.27	.34
	Positive	.33	.30	.25	.34	.33	.32
	Total	.25	.30	.23	.31	.28	.32
	Negative	.23	.36	.31	.43	.34	.29
Total	Neutral	.29	.29	.37	.39	.31	.32
	Positive	.36	.38	.30	.40	.32	.30
				Delay			
	Negative	.42	.27	.38	.25	.41	.24
V	Neutral	.40	.29	.47	.27	.37	.28
	Positive	.54	.25	.38	.23	.41	.26
	Total	.44	.27	.41	.25	.39	.26
	Negative	.31	.30	.31	.23	.37	.25
G	Neutral	.27	.25	.43	.30	.42	.24
	Positive	.38	.25	.36	.26	.34	.28
	Total	.31	.27	.37	.27	.37	.26
	Negative	.31	.30	.27	.24	.38	.28
VG	Neutral	.27	.25	.49	.30	.31	.28
	Positive	.38	.25	.38	.29	.39	.30
	Total	.31	.27	.39	.29	.36	.29
	Negative	.34	.44	.32	.42	.38	.26
Total	Neutral	.32	.38	.46	.39	.38	.26
	Positive	.43	.53	.37	.39	.37	.28

On the delayed test, unrelated distractor acceptance was higher in the V condition than in the G and VG conditions, $F(2,180) = 16.25$, $p < .05$. As for mood, there was a reliable effect on unrelated distractor acceptance on the immediate test, $F(2,114) = 3.51$, $p < .05$, but not on the delayed test, $F(2,90) = 1.89$, *ns*. This effect on the immediate test occurred because unrelated distractor acceptance was higher for positive moods compared to negative moods (acceptance for neutral moods did not differ reliably from either emotional mood). Unrelated distractor acceptance was also highest for positive moods on the delayed test, although that difference was not reliable. Because the unrelated distractor acceptance was affected by both mood and question, the data were bias-corrected using the unrelated distractor acceptance rates separated by question and mood. The following analyses were conducted on the bias-corrected values.

Analysis of Standard True and False Memory

In order to compare the present findings to the large body of research that has measured mood effects on standard measures of true and false memory, I began by isolating the data to the instructional condition that closely represents standard recognition tests. In those tests, subjects are simply asked to indicate with “yes” or “no” whether a given item has been presented. In the present study, I thus isolated the data to the V instructional condition, where subjects responded to the statement “I heard this word during the study phase.” Thus, I was able to determine whether the present results are consistent with prior research before examining differences among the instructional conditions. This method was implemented for all four experiments.

The design was a one-way ANOVA with valence (negative, neutral, positive) as the between-subjects factor, with verbatim acceptance of targets and CDs on the immediate and delayed tests as the dependent variables. Valence did not have a reliable effect on true memory on the immediate test, but false memory was higher for subjects in positive moods ($M = .45$)

compared to negative moods (*M* = .33), consistent with prior research (false memory for subjects in neutral moods did not reliably differ from the other two mood groups), $F(1,89) = 6.57$, $p < .05$.

On the delayed test, negative moods (*M* = .19) increased true memory compared to neutral (*M* = .12) and positive (*M* = .08) moods, which did not differ from each other, $F(1,89) = 5.23$, $p < .05$. Valence did not have an effect on false memory on the delayed test, $F(2,89) = 1.72$, *ns*. Note that there is no evidence that mood affects true memory in prior research.

Main ANOVA Results

Separate 3 (question: V, G, VG) x 3 (valence: negative, neutral, positive) ANOVAs were conducted on target and CD bias-corrected acceptance rates, separately for the immediate and delayed memory tests, with question as a within-subjects factor and mood valence as a between-subjects factor.

Immediate test. Turning to targets first, the main effect of question revealed that target acceptance was highest in the VG condition (*M* = .45), followed by the V (*M* = .27) condition and finally the G (*M* = .17) conditions, $F(2,228) = 32.81$, $p < .05$; all mean differences were reliable. There was also a marginal main effect of valence, with target acceptance being greater for the negative (*M* = .32) and neutral (*M* = .30) mood groups compared to the positive mood group (*M* = .24), $F(2,114) = 2.93$, $p = .06$.

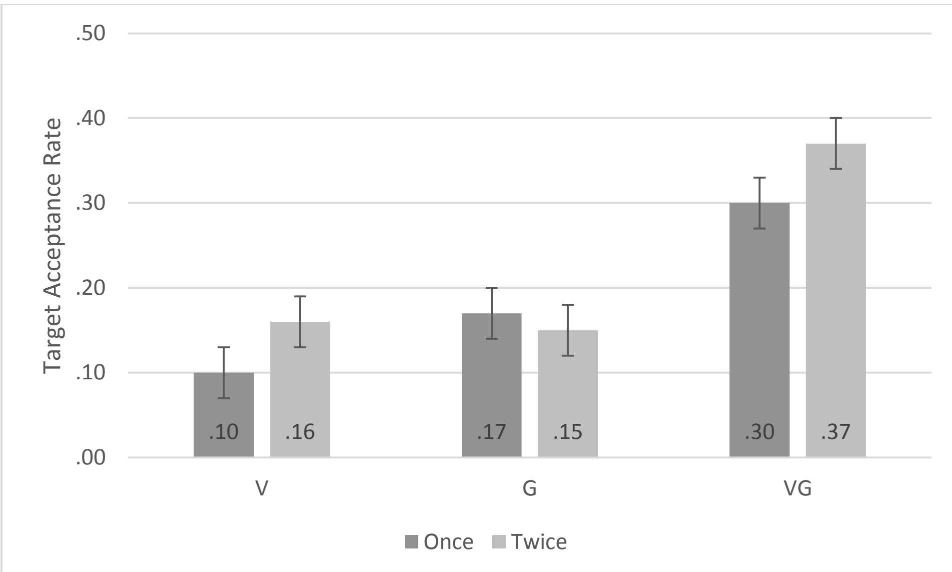
Moving on to CDs, on the immediate test the effect of question was reliable, but there was no main effect of valence or an interaction even though valence affected CD acceptance when isolated to the V condition. The main effect of question was such that CD acceptance was highest in the VG condition (*M* = .55), followed by the V condition (*M* = .41), then the G condition (*M* = .23); all mean differences were reliable, $F(2,228) = 27.41$, $p < .05$.

Delayed test. For targets there was again a reliable main effect of question, such that target acceptance was greater in the VG condition (M = .33) than the V (M = .13) and G (M = .16) conditions, which did not differ reliably from each other, $F(2,178) = 30.18$, $p < .05$. Note that verbatim memory declined over the delay period but gist memory did not. There was also a main effect of valence, with target acceptance being greater for the negative mood group (M = .25) compared to the positive mood group (M = .14); the neutral mood group (M = .22) did not differ reliably from the other two groups, $F(2,89) = 3.94$, $p < .05$.

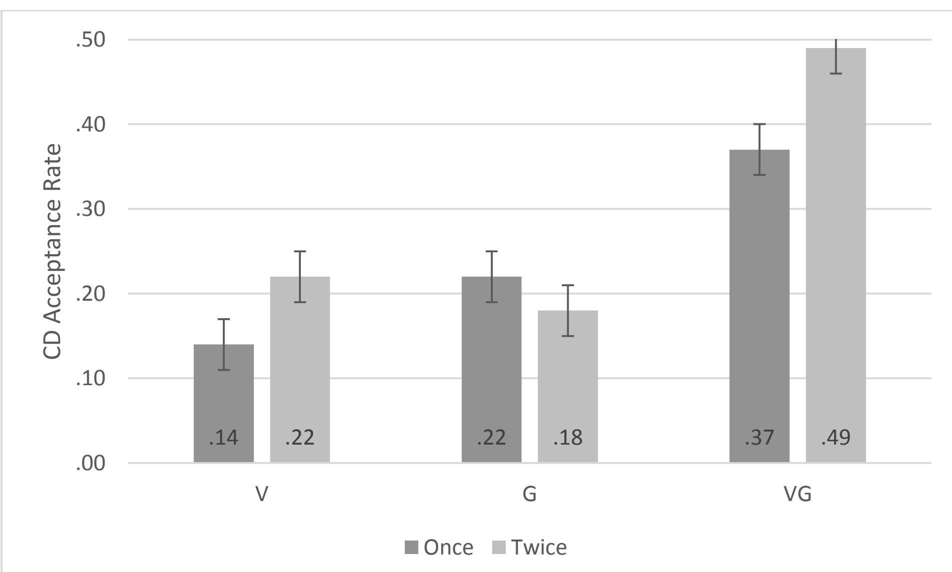
For critical distractors the main effect of question was similar to the effect for targets, such that CD acceptance was greater in the VG condition (M = .42) than the V (M = .18) and G (M = .20) conditions, which did not differ from each other, $F(2,178) = 39.92$, $p < .05$. There was no effect of mood valence on the delayed test.

Prior testing. Another set of ANOVAs was conducted in order to compare acceptance rates of targets and CDs on the delayed test that had not been tested on the immediate test to targets and CDs on the delayed test that had been tested on the immediate test. That is, this comparison was an analysis of the effect of repeated testing, and applies only to items on the delayed test. The design was thus two parallel 3 (question: V, G, VG) x 3 (valence: negative, neutral, positive) x 2 (times tested: once, twice) ANOVAs conducted on delayed test target and CD acceptance rates.

Turning to targets first, there was a main effect of prior testing (times tested), because items that were tested twice (M = .23) were more likely to be accepted than items that were tested only once (M = .19), $F(1,89) = 7.55$, $p < .05$. Furthermore, the question x prior testing interaction revealed that testing twice increased target acceptance in the V and VG conditions, but prior testing did not affect items in the G condition (see Figure 3.1, panel A), suggesting that this manipulation affects verbatim but not gist retrieval, $F(2,178) = 5.08$, $p < .05$. Finally, the valence x prior testing interaction revealed that prior testing increased target acceptance in the negative mood group (.22 versus .29) but not in the neutral (.22 versus .21) and positive (.12 versus .16) groups, $F(2,89) = 3.15$, $p = .05$. This finding, considered with the previous effect, is consistent with the proposition that negative moods enhance verbatim retrieval.



A



B

Figure 3.1. Effect of prior testing and question on delayed test targets (panel A) and CDs (Panel B) in Experiment 1.

Turning now to CDs, again there was a reliable effect of prior testing, because CDs that were tested twice (M = .30) were more likely to be accepted than CDs that were only tested once (M = .24), $F(1,89) = 6.64$, $p < .05$. There was also a reliable question x prior testing interaction that was similar to the parallel interaction for targets, $F(2,178) = 6.89$, $p < .05$. Prior testing had no effect on CDs in the G condition (although mean was higher for *un*tested items compared to tested items), but CDs in the V and VG conditions were more likely to be accepted when they were presented twice as compared to those presented only once (see Figure 3.1, panel B).

Forgetting. In order to measure subjects' forgetting over the 48-hour delay period, additional ANOVAs were computed with test (immediate vs. delayed) as a factor, excluding items on the delayed test that were being tested for a second time. That is, a comparison was made between items that were tested immediately after the study phase and items tested for the first time two days after the study phase. The design was thus two parallel 3 (question: V, G, VG) x 3 (valence: negative, neutral, positive) x 2 (test: immediate, delayed) ANOVAs with target and CD acceptance rates as the dependent variables.

Turning to targets first, acceptance was higher on the immediate test (M = .31) than on the delayed test (M = .19), $F(1,90) = 18.18$, $p < .05$. Furthermore, the test x question interaction supported the prediction that verbatim memory traces would decline more rapidly over time, $F(2,180) = 11.65$, $p < .05$ (see Figure 3.2, panel A): in the V and VG condition, target acceptance was higher on the immediate test than on the delayed test but in the G condition target acceptance was equivalent across tests. That is, forgetting occurred for verbatim acceptance but not gist acceptance of targets.

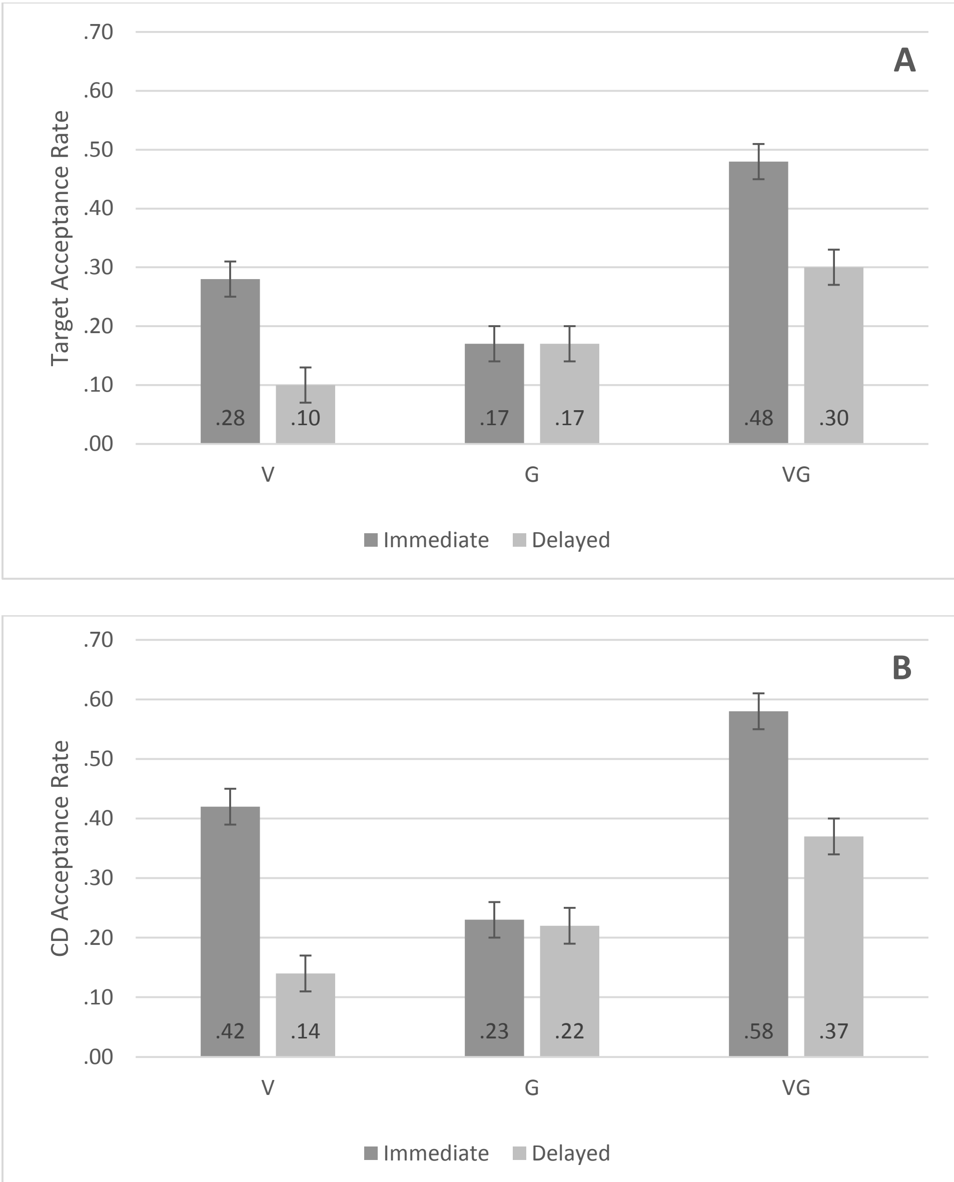


Figure 3.2. Forgetting of targets (panel A) and CDs (panel B) by question (Experiment 1).

Similarly, for critical distractors, the main effect of test and the question x test interaction were reliable. The main effect of test was similar to the effect on targets because the critical distractor acceptance rate decreased from the immediate (M = .41) to the delayed test (M = .25), $F(1,90) = 28.18$, $p < .05$. The question x test interaction, which was also similar to the parallel interaction for targets, revealed that CD acceptance in the V and VG conditions declined over the delay while CD acceptance in the G remained stable over the delay (see Figure 3.2, panel B), $F(2,180) = 15.92$, $p < .05$.

Summary. Examination of standard true and false memory revealed two results: a) negative moods reduced false memory compared to positive moods immediately after list presentation and b) negative moods increased true memory compared to positive moods after a two-day delay. Both of these effects are consistent with the prediction that mood effects on memory are driven by enhanced item-specific processing, or enhanced verbatim retrieval, when subjects are in negative moods. When the analysis was expanded to all three instructional conditions, negative moods increased target acceptance on the immediate test as well.

The second major set of results concerned prior testing. In that connection, testing items twice a) increased both target and CD acceptance compared to testing items only once and b) increased target and CD acceptance in the V condition but not the G condition. These results are consistent with the idea that repeated testing affects verbatim retrieval. Furthermore, prior testing increased target acceptance for subjects in negative moods but not positive and neutral moods, consistent with the idea that negative moods increase reliance on verbatim retrieval.

Finally, forgetting occurred for both targets and CDs, and affected responses in the V and VG conditions but not the G condition. Thus, similar to repeated testing, forgetting impacts verbatim memory traces more than gist traces, because gist traces remain more stable over time. We next turn to the results of the conjoint recognition model to determine whether the parameter estimates provide further support for the hypothesis that negative mood, forgetting, and repeated testing all impact verbatim processes more than gist processes.

Model Fits

The model's fit was evaluated by computing likelihood ratio tests, which generate G^2 statistics asymptotically distributed with χ^2 with one degree of freedom per condition. The critical value for rejecting the null hypothesis that the model fits the data is 3.84. The G^2 values for all of the conditions were below 3.84, indicating that the model provided a statistically acceptable estimation of the data.

Parameter Estimates

Regarding the first of the above effects, recall that the identity judgment parameter is a measure of a verbatim process which allows subjects to correctly accept targets on the basis of matching verbatim traces. An examination of the parameter estimates in Table 3.4 confirms that there was variability in identity judgment as a result of valence. Recall that negative moods increased target acceptance compared to positive moods on both tests, and reduced false memory on the immediate test. Consistent with the first effect, identity judgment was .29 for negative mood and .16 for positive mood on the immediate test, and .19 for negative mood and .11 for positive mood on the delayed test. However, there were no reliable differences in the false memory parameters on the immediate test as a function of valence.

Turning now to the second major ANOVA result, concerning prior testing, testing items twice increased verbatim acceptance of targets and CDs. For targets, testing items twice increased acceptance for subjects in negative moods only, and, consistently, identity judgment increased with prior testing for negative moods but not positive and neutral moods. On the false memory side, familiarity and phantom recollection were both higher for items that had been previously tested, suggesting that despite being driven by gist processes, the prior testing effect

Table 3.4

Experiment 1 Parameter Estimates

<table>
<tr><th></th><th colspan="2">R_T</th><th colspan="2">F_T</th><th colspan="2">P_D</th><th colspan="2">R_D</th><th colspan="2">F_D</th><th colspan="2">β_V</th><th colspan="2">β_G</th><th colspan="2">β_{VG}</th></tr>
<tr><td colspan="17">Immediate Test</td></tr>
<tr><td>Negative</td><td colspan="2">.29</td><td colspan="2">.14</td><td colspan="2">.35</td><td colspan="2">.12</td><td colspan="2">.32</td><td colspan="2">.34</td><td colspan="2">.34</td><td colspan="2">.30</td></tr>
<tr><td>Neutral</td><td colspan="2">.22</td><td colspan="2">.13</td><td colspan="2">.34</td><td colspan="2">.15</td><td colspan="2">.33</td><td colspan="2">.39</td><td colspan="2">.35</td><td colspan="2">.39</td></tr>
<tr><td>Positive</td><td colspan="2">.16</td><td colspan="2">.13</td><td colspan="2">.31</td><td colspan="2">.16</td><td colspan="2">.29</td><td colspan="2">.38</td><td colspan="2">.33</td><td colspan="2">.38</td></tr>
<tr><td>Mean</td><td colspan="2">.24</td><td colspan="2">.12</td><td colspan="2">.34</td><td colspan="2">.15</td><td colspan="2">.31</td><td colspan="2">.37</td><td colspan="2">.34</td><td colspan="2">.34</td></tr>
<tr><td colspan="17">Delayed Test</td></tr>
<tr><td></td><td>U</td><td>T</td><td>U</td><td>T</td><td>U</td><td>T</td><td>U</td><td>T</td><td>U</td><td>T</td><td>U</td><td>T</td><td>U</td><td>T</td><td>U</td><td>T</td></tr>
<tr><td>Negative</td><td>.16</td><td>.25</td><td>.17</td><td>.20</td><td>.15</td><td>.41</td><td>.03</td><td>.14</td><td>.28</td><td>.33</td><td>.37</td><td>.44</td><td>.32</td><td>.29</td><td>.32</td><td>.44</td></tr>
<tr><td>Neutral</td><td>.18</td><td>.14</td><td>.11</td><td>.12</td><td>.21</td><td>.25</td><td>.17</td><td>.11</td><td>.20</td><td>.39</td><td>.46</td><td>.48</td><td>.35</td><td>.28</td><td>.35</td><td>.48</td></tr>
<tr><td>Positive</td><td>.10</td><td>.14</td><td>.13</td><td>.07</td><td>.14</td><td>.40</td><td>.24</td><td>.12</td><td>.13</td><td>.17</td><td>.48</td><td>.46</td><td>.23</td><td>.40</td><td>.39</td><td>.53</td></tr>
<tr><td>Mean</td><td>.16</td><td>.19</td><td>.15</td><td>.10</td><td>.21</td><td>.37</td><td>.16</td><td>.14</td><td>.19</td><td>.29</td><td>.44</td><td>.47</td><td>.32</td><td>.33</td><td>.32</td><td>.47</td></tr>
<tr><td>Negative</td><td colspan="2">.19</td><td colspan="2">.17</td><td colspan="2">.24</td><td colspan="2">.07</td><td colspan="2">.30</td><td colspan="2">.40</td><td colspan="2">.30</td><td colspan="2">.40</td></tr>
<tr><td>Neutral</td><td colspan="2">.12</td><td colspan="2">.14</td><td colspan="2">.19</td><td colspan="2">.11</td><td colspan="2">.31</td><td colspan="2">.45</td><td colspan="2">.30</td><td colspan="2">.45</td></tr>
<tr><td>Positive</td><td colspan="2">.11</td><td colspan="2">.09</td><td colspan="2">.29</td><td colspan="2">.17</td><td colspan="2">.14</td><td colspan="2">.46</td><td colspan="2">.32</td><td colspan="2">.46</td></tr>
<tr><td>Mean</td><td colspan="2">.14</td><td colspan="2">.14</td><td colspan="2">.24</td><td colspan="2">.12</td><td colspan="2">.25</td><td colspan="2">.43</td><td colspan="2">.31</td><td colspan="2">.43</td></tr>
<tr><td colspan="17">Combined Across Both Tests</td></tr>
<tr><td>Negative</td><td colspan="2">.26</td><td colspan="2">.15</td><td colspan="2">.34</td><td colspan="2">.16</td><td colspan="2">.24</td><td colspan="2">.37</td><td colspan="2">.33</td><td colspan="2">.33</td></tr>
<tr><td>Neutral</td><td colspan="2">.18</td><td colspan="2">.14</td><td colspan="2">.28</td><td colspan="2">.13</td><td colspan="2">.32</td><td colspan="2">.41</td><td colspan="2">.35</td><td colspan="2">.41</td></tr>
<tr><td>Positive</td><td colspan="2">.14</td><td colspan="2">.11</td><td colspan="2">.30</td><td colspan="2">.16</td><td colspan="2">.23</td><td colspan="2">.41</td><td colspan="2">.33</td><td colspan="2">.41</td></tr>
<tr><td>Mean</td><td colspan="2">.19</td><td colspan="2">.14</td><td colspan="2">.30</td><td colspan="2">.13</td><td colspan="2">.30</td><td colspan="2">.39</td><td colspan="2">.33</td><td colspan="2">.39</td></tr>
</table>

on CDs results in acceptance in the V condition; i.e., false memory increases when CDs have been previously tested.

Finally, addressing the forgetting results, I compared the parameter values for the immediate test to previously untested delayed test items. Recall here that target and CD acceptance declined in the V condition but remained stable in the G condition. Consistently, estimates of identity judgment declined from the immediate test to the delayed test, whereas target familiarity remained stable over the delay. For CDs the only parameter that declined over time was phantom recollection, again suggesting that this parameter drove CD acceptance in the V condition.

Discussion

Experiment 1 provided preliminary results regarding the effect of low-arousal moods on true and false memory. This study was the first to control arousal at a low level while manipulating valence in a false memory design, and was the first to provide information on retrieval processes associated with different mood valences as well as forgetting and testing effects. As such, there is a dearth of previous research with which to compare the results.

I found that negative mood valence increased true memory by increasing verbatim retrieval, which is consistent with the prediction that negative mood results in more verbatim, item-specific processing. Past researchers have not provided evidence of mood effects on true memory, but I argue that this difference is due to the conflation of valence and arousal in previous studies. Positive and negative valence differences may become attenuated when arousal is high; valence effects may only be detectable when arousal is low, as in the present study. This could be the case if arousal increases verbatim processing. If low-arousal negative moods enhance true memory compared to low-arousal positive moods via verbatim processing, when verbatim processing is increased in positive moods due to increased arousal the valence effect may disappear. In Experiment 3 I investigate how valence effects change when arousal is increased, and in Experiment 4 I investigate whether it is true that arousal increases verbatim processing.

Furthermore, the effect of mood on true memory was reliable on the delayed test but not on the immediate test (although mood did affect target acceptance across all instructional conditions on the immediate test). If mood affects true memory (i.e., acceptance in the V condition) only after a delay, the lack of evidence for such an effect in prior research could be because the typical recall or recognition test occurs on the same day as the study phase. Negative mood reduced false memory compared to positive mood on the immediate test; however, the only reliable parameter differences were in the identity parameter. The results were consistent with both ideas that negative mood would reduce false memory and that it would increase verbatim processing, although the data did not indicate any parameter differences on the false memory side, per se. The effect of negative mood was consistent with findings of the same nature, including those lacking control over arousal (e.g., Storbeck & Clore, 2005, 2011).

Experiment 1 also provided further evidence regarding manipulations that primarily affect verbatim processing: namely, repeated testing and forgetting. Repeated testing primarily affected the identity parameter, and forgetting resulted in declines in identity and phantom recollection. In addition to providing process explanations for repeated testing and forgetting effects, these results provide further confirmation of the prediction from FTT that verbatim memory traces decline more rapidly over time than gist traces.

It is possible that I did not detect false memory effects on the delayed test because the false acceptance rate was simply too low (CD acceptance in the V condition was .18 for

Experiment 1, as compared to .24 in Experiment 2 and .30 in Experiment 3). If false memory was at floor level, perhaps subjects in all three mood conditions were able to successfully reject CDs, and there was not enough latitude for variation in false memory rate as a function of mood. In the next study, I included a manipulation that would differentially affect true memory and false memory as well as provide further testing of the verbatim processing in negative mood hypothesis.

CHAPTER 4

EXPERIMENT 2: EFFECT OF MOOD VALENCE AND LIST LENGTH ON TRUE AND FALSE MEMORY FOR NEUTRAL WORDS

Using the same general design and the same mood manipulation as Experiment 1, in Experiment 2 I introduced a list length manipulation designed to strengthen verbatim memory traces (in one list length) and promote more false memories (in another list length). Prior research has shown that a) shortening DRM lists to as few as just one item makes it easier to use verbatim retrieval (Brainerd & Reyna, 2007) and b) lengthening lists up to around 14 items increases false memory rates (Robinson & Roediger, 1997). In place of the 10-item lists used in Experiment 1, Experiment 2 used a) 6-item lists to promote verbatim retrieval and b) 14-item DRM lists in order to heighten false memory rates as much as possible and thus eliminate the possibility that false memory rates were simply too low in Experiment 1. List length has been demonstrated to be a manipulation that dissociates verbatim and gist processing (Sugrue & Hayne, 2006), so if mood interacts with list length it could provide further information regarding the contribution of the two types of processing to different moods.

Method

Participants

112 subjects participated in exchange for course credit. 94 subjects were women and all were between the ages of 18 and 24 (M_{age} = 19.61, *SD* = 1.14). 33 subjects were in the negative condition, 39 were in the neutral condition, and 40 were in the positive condition.

Materials

The mood induction materials were identical to those used in Pilot 2. They included the pre- and post-induction mood measurements and the three videos designed to induce positive, negative, and neutral moods with equivalent low levels of arousal.

The memory task materials consisted of 234 words, selected in the same manner as done in Experiment 1. However, rather than 18 10-item lists, 9 6-item lists and 9 14-item lists were used (see Table 4.1). The 6-item lists were simply the first 6 targets of those lists in Experiment 1, and the 14-item lists were comprised of the 10 items used in Experiment 1 plus the subsequent 4 targets from Stadler et al.'s (1999) 15-item lists. The remaining target was used as the RD. The materials were 18 CDs, 18 RDs, 18 URDs, and 180 targets, as in Experiment 1. New audio files were created for presentation of the 18 lists. The words were read aloud in order of descending associative strength at a two-second rate.

Table 4.2 displays how the conjoint recognition test questions were assigned to the items on the immediate and delayed test. The breakdown was similar to the breakdown of Experiment 1 items, with the further qualification that on the delayed test, 9 of the CDs, 9 of the RDs, and 18 of the targets were from 6-item (short) lists and the other half of the items were from 14-item (long lists.) As in Experiment 1, only half of the items were tested on the immediate test and CR questions were distributed equally among items. The immediate test items were re-tested on the delayed test along with the other, untested half of items, and the re-tested items were accompanied by the same CR questions as on the immediate test. As in Experiment 1, 72 items were on the immediate test and 144 items were on the delayed test.

Table 4.1

Experiment 2 and 3 Items

CD	*Doctor*	*needle*	*sleep*	*smoke*	*foot*	*sweet*	*trash*	*mountain*	*cup*
T1	Nurse	thread	bed	cigarette	toe	sour	garbage	hill	saucer
T2	physician	syringe	nap	cigar	shoe	candy	dump	climber	mug
T3	Stethoscope	pin	rest	pipe	inch	sugar	can	climb	glass
T4	Surgeon	haystack	tired	fire	ankle	bitter	litter	peak	measuring
T5	Patient	sewing	snooze	chimney	hand	honey	junk	valley	coffee
T6	Dentist	injection	drowsy	puff	leg	tart	waste	summit	tea
T7			slumber			chocolate	landfill		soup
T8			dream			taste	bag		drink
T9			snore			tooth	rubbish		straw
T10			wake			heart	debris		sip
T11			blanket			good	refuse		lid
T12			doze			soda	sewage		handle
T13			peace			cake	sweep		goblet
T14			awake			pie	scraps		plastic
RD	Cure	sharp	yawn	tobacco	knee	nice	pile	top	coaster
UD	Shirt	lion	thief	music	army	rubber	king	car	fruit

CD	*River*	*city*	*chair*	*cold*	*window*	*smell*	*soft*	*rough*	*slow*
T1	Stream	town	table	hot	pane	aroma	hard	smooth	fast
T2	Creek	suburb	seat	chilly	sill	scent	pillow	sandpaper	snail
T3	Flow	state	couch	shiver	shutter	stench	gentle	tough	turtle
T4	Lake	county	sit	arctic	glass	sniff	smooth	bumpy	quick
T5	Brook	country	stool	frigid	door	fragrance	cotton	uneven	molasses
T6	Water	street	desk	ice	curtain	nose	skin	coarse	speed
T7				freeze	ledge	hear	feather	gravel	
T8				warm	shade	see	plush	ground	
T9				frost	view	nostril	loud	jagged	
T10				heat	open	breathe	tender	sand	
T11				snow	sill	reek	light	road	
T12				wet	frame	whiff	fur	ready	
T13				weather	breeze	rose	fluffy	boards	
T14				air	screen	odor	downy	riders	
RD	Bridge	big	recliner	winter	house	stink	touch	rugged	stop
UD	Black	girl	pen	flag	angry	spider	high	man	bread

Table 4.2

Experiment 2 and 3 Test Items with CR Statement

Type	*Word*	*Times Tested*	*CR*	*Length*	*Type*	*Word*	*Times Tested*	*CR*	*Length*
CD	window	1	G	14	CD	chair	2	G	6
CD	mountain	1	G	6	CD	rough	2	G	14
CD	cup	1	G	14	CD	slow	2	G	6
CD	needle	1	V	6	CD	river	2	V	6
CD	sweet	1	V	14	CD	cold	2	V	14
CD	trash	1	V	14	CD	city	2	V	6
CD	doctor	1	VG	6	CD	smell	2	VG	14
CD	sleep	1	VG	14	CD	foot	2	VG	6
CD	smoke	1	VG	6	CD	soft	2	VG	14
RD	cure	1	G	6	RD	coaster	2	G	14
RD	sharp	1	G	6	RD	bridge	2	G	6
RD	nice	1	G	14	RD	rugged	2	G	14
RD	knee	1	V	6	RD	recliner	2	V	6
RD	touch	1	V	14	RD	stink	2	V	14
RD	big	1	V	6	RD	winter	2	V	14
RD	yawn	1	VG	14	RD	stop	2	VG	6
RD	tobacco	1	VG	6	RD	house	2	VG	14
RD	pile	1	VG	14	RD	top	2	VG	6
T1	injection	1	G	6	T1	street	2	G	6
T1	drowsy	1	G	14	T1	desk	2	G	6
T1	Tea	1	G	14	T1	coarse	2	G	14
T1	dentist	1	V	6	T1	curtain	2	V	14
T1	tart	1	V	14	T1	skin	2	V	14
T1	water	1	V	6	T1	speed	2	V	6
T1	leg	1	VG	6	T1	summit	2	VG	6
T1	puff	1	VG	6	T1	ice	2	VG	14
T1	waste	1	VG	14	T1	nose	2	VG	14
T2	physician	1	G	6	T2	climber	2	G	6
T2	candy	1	G	14	T2	scent	2	G	14
T2	suburb	1	G	6	T2	sandpaper	2	G	14
T2	nap	1	V	14	T2	mug	2	V	14
T2	shoe	1	V	6	T2	seat	2	V	6
T2	dump	1	V	14	T2	snail	2	V	6
T2	syringe	1	VG	6	T2	creek	2	VG	6
T2	cigar	1	VG	6	T2	sill	2	VG	14
T2	chilly	1	VG	14	T2	pillow	2	VG	14
UD	shirt	1	G		UD	pen	2	G	

UD	music	1	G	UD	angry	2	G
UD	fruit	1	G	UD	man	2	G
UD	rubber	1	V	UD	spider	2	V
UD	king	1	V	UD	high	2	V
UD	car	1	V	UD	bread	2	V
UD	lion	1	VG	UD	black	2	VG
UD	thief	1	VG	UD	girl	2	VG
UD	army	1	VG	UD	flag	2	VG

Procedure

The procedure was identical to Experiment 1.

Results

Mood manipulation check

Participants whose mood was not manipulated as intended were excluded from the analyses, resulting in 110 subjects (33 in the negative mood group, 39 in the neutral mood group, and 38 in the positive mood group). Participants were excluded if the change in their valence rating from the pre-test to the post-test was more than 2 points in the direction opposite to the intended (e.g., participants in the positive group whose moods became substantially more negative), and participants were excluded if the change in their arousal rating from the pre-test to the post-test was more than two points (i.e., if there was a substantial increase in arousal). Analysis of the resulting 110 subjects revealed reliable differences in post-test valence for the three groups, $F(2,109) = 347.34$, $p < .01$, with valence being the highest for the positive group ($M = 7.32$, $SD = .53$), followed by the neutral group ($M = 5.59$, $SD = .55$), followed by the negative group ($M = 3.21$, $SD = .86$). Recall that these values are equivalent to those found in the previous study. As expected, there were no significant differences in arousal among the three groups, $F(2,109) = .86$, *ns*. ($M_{positive} = 4.88$, $M_{neutral} = 4.31$, $M_{negative} = 4.09$).

Bias correction

The effects of conjoint recognition question and mood on unrelated distractor acceptance were identified in order to determine the proper method of bias correction. Two parallel 3 (question: V, G, VG) x 3 (mood: positive, negative, neutral) analyses of variance (ANOVAs) were performed on the acceptance rates for unrelated distractors, separately for the immediate and the delayed tests. There was a reliable main effect of question on the immediate test but not the delayed test (see Table 3.3 for descriptive statistics). On the immediate test, unrelated distractor acceptance was higher in the G condition than the V condition, which was higher than the VG condition, $F(2,214) = 19.20$, $p < .05$. On the delayed test the effect of question on unrelated distractors was not reliable, $F(2,172) = 1.07$, *ns*. Because the unrelated distractor acceptance was affected by question, and because the question effect was not consistent over the delay period, the data were bias-corrected using the unrelated distractor acceptance rates separated by question and test. All of the following analyses were conducted on the bias-corrected values.

Analysis of Standard True and False Memory

As in Experiment 1, I began by investigating the effect of valence—and now the additional effect of list length—on standard true and false memory. Valence did not affect true memory on the immediate test, but it did have a marginal effect on false memory because negative moods ($M = .29$) reduced false memory compared to positive moods ($M = .40$), as in Experiment 1, $F(2,107) = 4.12$, $p = .05$. Additionally, longer lists ($M = .32$) increased false memory compared to short lists ($M = .24$) on the delayed test, $F(1,86) = 8.44$, $p < .05$. There were no effects on true memory and there were no mood effects on false memory on the delayed test.

I also analyzed the effect of forgetting on true and false memory, by comparing true and false memory on the immediate test to true and false memory for delayed test items that had not been previously tested. Forgetting affected true memory because it was higher on the immediate test (.31) than the delayed test (.13), $F(1,86) = 44.93$, $p < .05$. Forgetting occurred for false memory as well, because it was higher on the immediate test ($M = .34$) than the delayed test ($M = .27$), $F(1,86) = 4.98$, $p < .05$. This effect was driven by short lists, because long lists were remembered equivalently on both tests (.36 vs. .35) whereas false memory for short lists was higher on the immediate test ($M = .32$) compared to the delayed test ($M = .16$), $F(1,86) = 5.38$, $p < .05$.

Main ANOVA Results

Separate 3 (question: V, G, VG) x 3 (valence: negative, neutral, positive) x 2 (list length: short, long) ANOVAs were conducted on target and CD bias-corrected acceptance rates, separately for the immediate and delayed memory tests, with question and length as within-subjects factors and valence as the between-subjects factor.

Immediate test. Turning to targets first, the main effect of question revealed that target acceptance was highest in the VG condition ($M = .48$), followed by the V condition ($M = .32$), then the G condition ($M = .13$); all mean differences were reliable, $F(2,214) = 34.41$, $p < .05$. There was no main effect of valence, even though this effect was reliable in Experiment 1.

Moving on to CDs, on the immediate test the effect of question was reliable, as was the question x length interaction. None of the other main effects or interactions were reliable. The main effect of question was similar to the effect for targets: CD acceptance was higher in the VG condition ($M = .59$) than the V condition ($M = .34$) and the G condition ($M = .29$), $F(2,214) = 18.68$, $p < .05$. The question x length interaction (see Figure 4.1) revealed that CD acceptance was higher for shorter lists compared to longer lists in the G condition, but the difference was not reliable in the other two conditions, $F(2,214) = 4.34$, $p < .05$.

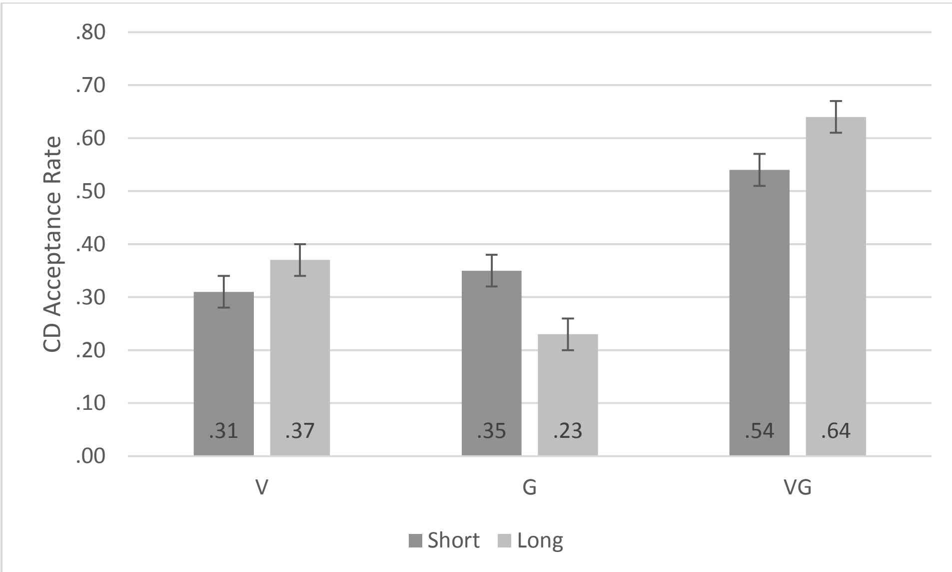


Figure 4.1. CD acceptance rate by question and length on the immediate test (Experiment 2).

Delayed test. For targets the effect of question was such that target acceptance was greater in the VG condition ($M = .33$) than the V ($M = .19$) and G ($M = .18$) conditions, which did not differ, $F(2,172) = 18.10$, $p < .05$. There was also a main effect of length, with target acceptance being higher for short lists ($M = .28$) than long lists ($M = .24$), $F(1,86) = 14.28$, $p < .05$. Finally, there was a length x question interaction such that shorter lists increased target acceptance in the G condition (.31 versus .20) but not the V or VG condition, $F(2,172) = 6.94$, $p < .05$.

For CDs there was a reliable main effect of question and a reliable main effect of length. The effect of question was such that CD acceptance was greater in the VG condition ($M = .37$) than the V (.28) and G conditions ($M = .31$), which did not differ, $F(2,172) = 9.89$, $p < .05$. The main effect of length occurred because CD acceptance was higher for longer lists (.34) than shorter lists (.30), $F(1,86) = 7.78$, $p < .05$.

Prior testing. Another set of ANOVAs were conducted in order to compare acceptance rates of targets and critical distractors on the delayed test that had not been tested on the immediate test to targets and critical distractors on the delayed test that had been tested on the immediate test. That is, this comparison was an analysis of the effect of prior testing, and applies only to items on the delayed test. The design was thus two parallel 3 (question: V, G, VG) x 3 (valence: negative, neutral, positive) x 2 (length: short, long) x 2 (times tested: once, twice) ANOVAs conducted on delayed test target and CD acceptance rates.

Turning first to targets, there was a reliable interaction between prior testing and question, prior testing and length, as well as a reliable three-way interaction between prior testing, question, and length. The question x prior testing interaction revealed that testing targets twice increased target acceptance in the V condition but reduced target acceptance in the G condition (see Figure 4.2), $F(1,172) = 7.75$, $p < .05$.

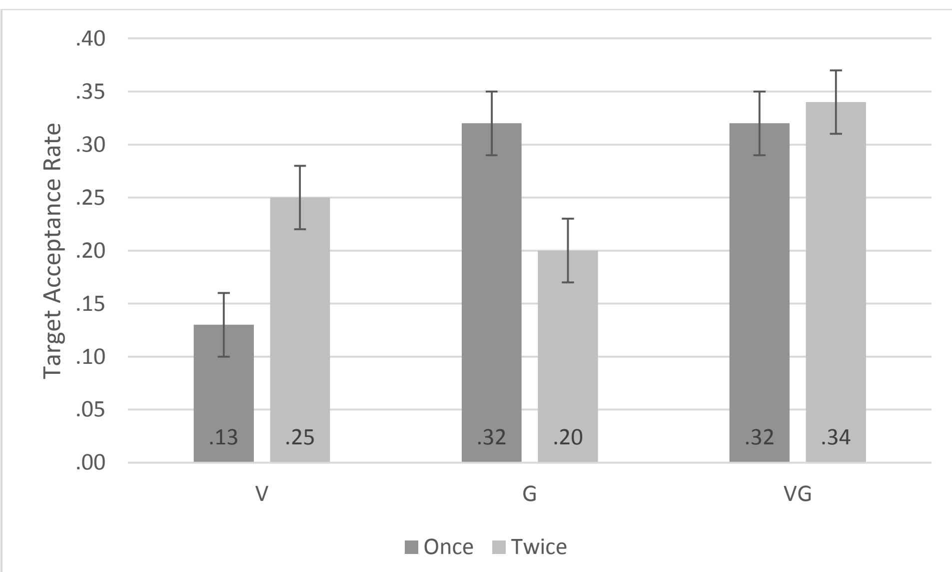


Figure 4.2. Target acceptance rate by prior testing and question on the delayed test (Experiment 2).

The length x prior testing interaction revealed that prior testing reduced target acceptance for short lists (.31 vs. .25) but increased target acceptance for long lists (.21 vs. .27), $F(1,86) = 27.28$, $p < .05$. The three-way interaction, shown in Figure 4.3, indicated that testing items twice increased target acceptance in the V condition for both list lengths, but reduced target acceptance in the G condition for short lists only, $F(2,172) = 4.70$, $p < .05$.

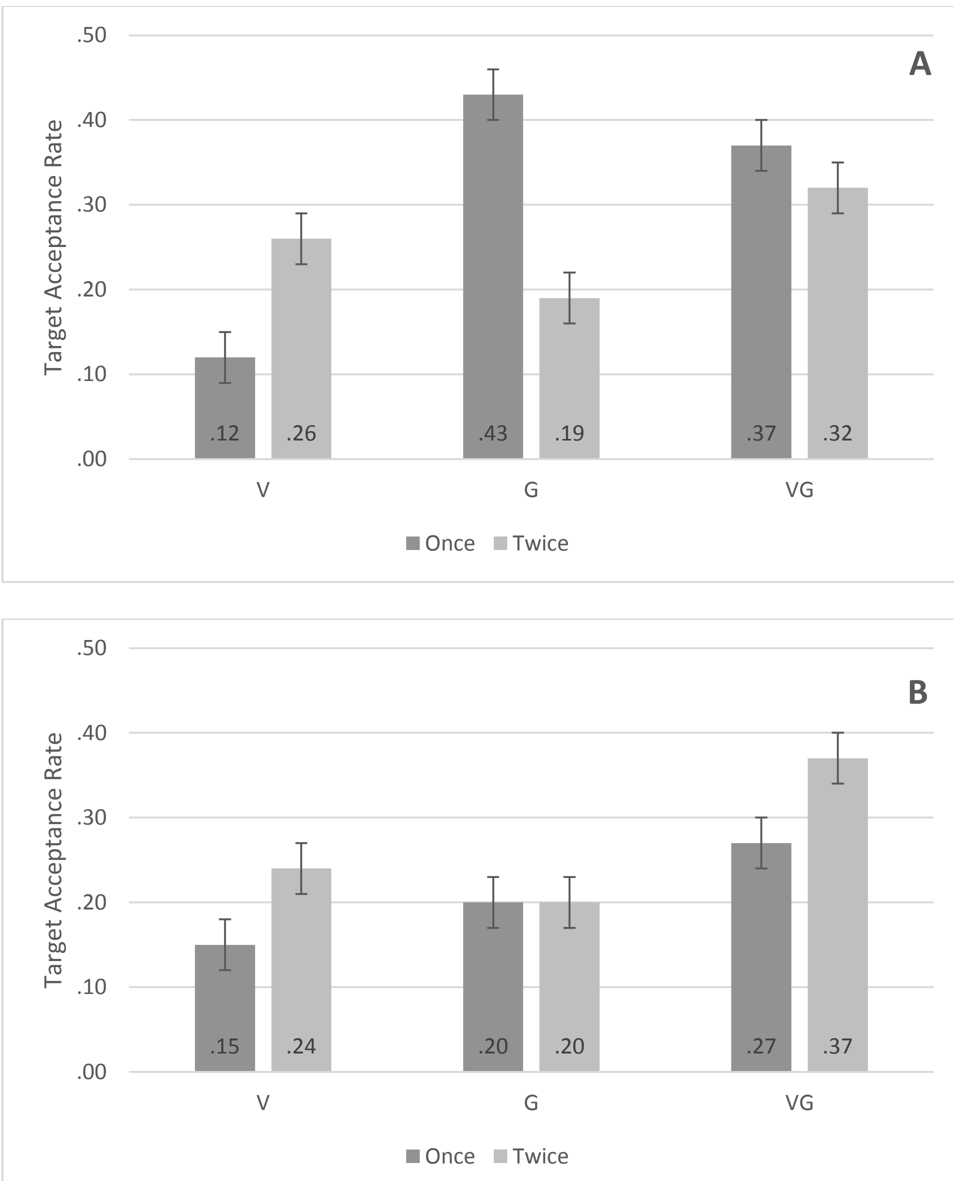


Figure 4.3. Target acceptance rate by prior testing and question for short lists (panel A) and long lists (panel B) on the delayed test (Experiment 2).

Turning to CDs, the main effect of prior testing was reliable because prior testing increased CD acceptance (.29 versus .35), $F(1,86) = 8.39$, $p < .05$. The question x prior testing and length x prior testing interactions were both reliable, as was the question x length x prior testing interaction. As for question x prior testing, testing items twice reduced CD acceptance in the G condition and increased CD acceptance in the VG condition, but did not affect CD acceptance in the V condition, $F(2,172) = 13.20$, $p < .05$. In terms of length, testing items twice increased CD acceptance for short lists but did not affect CD acceptance for long lists, $F(2,172) = 37.83$, $p < .05$. The three-way interaction, shown in Figure 4.4, revealed that prior testing increased CD acceptance in all conditions for short lists, but reduced CD acceptance in the G condition for long lists, $F(2,172) = 6.38$, $p < .05$.

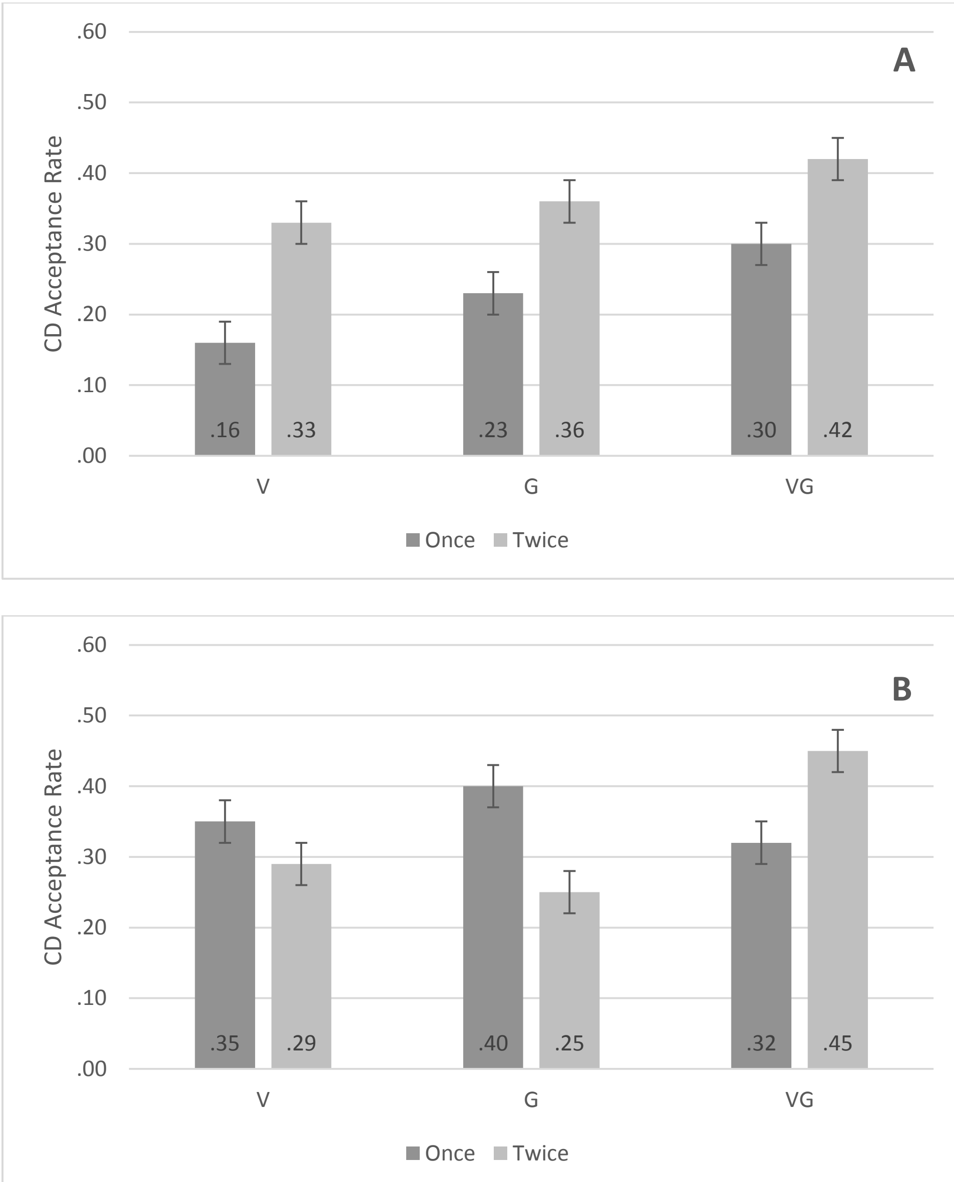


Figure 4.4. CD acceptance rate by question and list length for short lists (panel A) and long lists (panel B) on the delayed test (Experiment 2).

Forgetting. In order to measure subjects' forgetting over the two-day delay period, additional ANOVAs were computed with test (immediate vs. delayed) as a factor, excluding items on the delayed test that were being tested for a second time. That is, a comparison was made between items that were tested immediately after the study phase and items tested for the first time two days after the study phase. The design was two parallel 3 (question: V, G, VG) x 3 (valence: negative, neutral, positive) x 2 (length: short, long) x 2 (test: immediate, delayed) ANOVAs conducted on target and CD acceptance rates.

Turning to targets first, the ANOVA yielded a reliable main effect of test, a test x question interaction, a test x length interaction, and a reliable three-way interaction between question, length, and test. The target acceptance rate was higher on the immediate test (M = .31) than on the delayed test (M = .26), $F(1,86) = 12.30$, $p < .05$. Note that this pattern the same as the pattern found in Experiment 1. Furthermore, the test x question interaction again supported the prediction that verbatim memory traces would decline more rapidly over time, $F(2,172) = 18.02$, $p < .05$ (see Figure 4.5): in the V (.31 vs. .13) and VG (.48 vs. .32) conditions, target acceptance was higher on the immediate test than on the delayed test but in the G condition target acceptance was lower on the immediate test (.18) than on the delayed test (.32). Forgetting did not impact short lists, but target acceptance for long lists declined over time (.32 versus .21), $F(1,86) = 16.31$, $p < .05$.

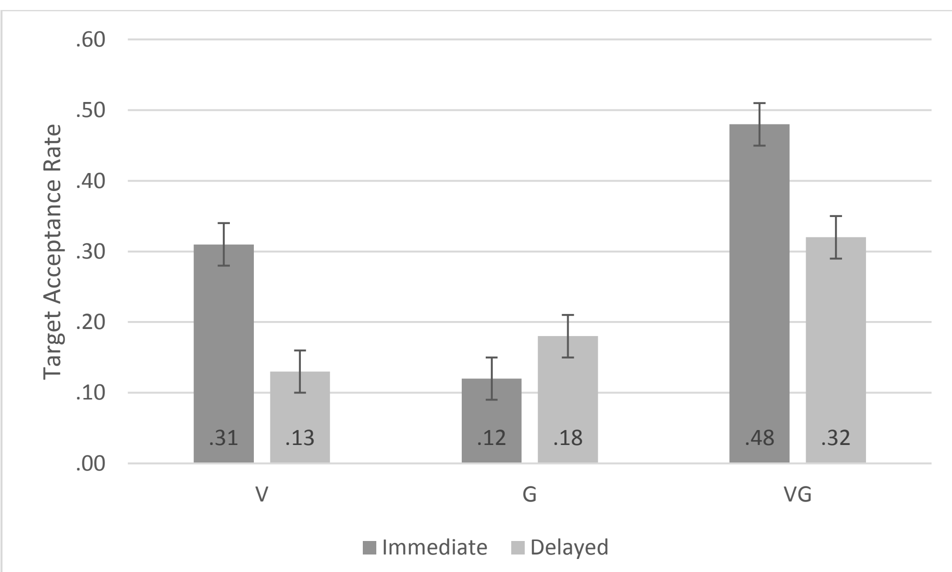


Figure 4.5. Forgetting of targets by question (Experiment 2).

Finally, the three-way interaction, shown in Figure 4.6, revealed that forgetting reduced target acceptance in the V condition for both list lengths, but target acceptance in the G condition increased over time for short lists but not long lists, $F(2,172) = 5.31$, $p < .05$.

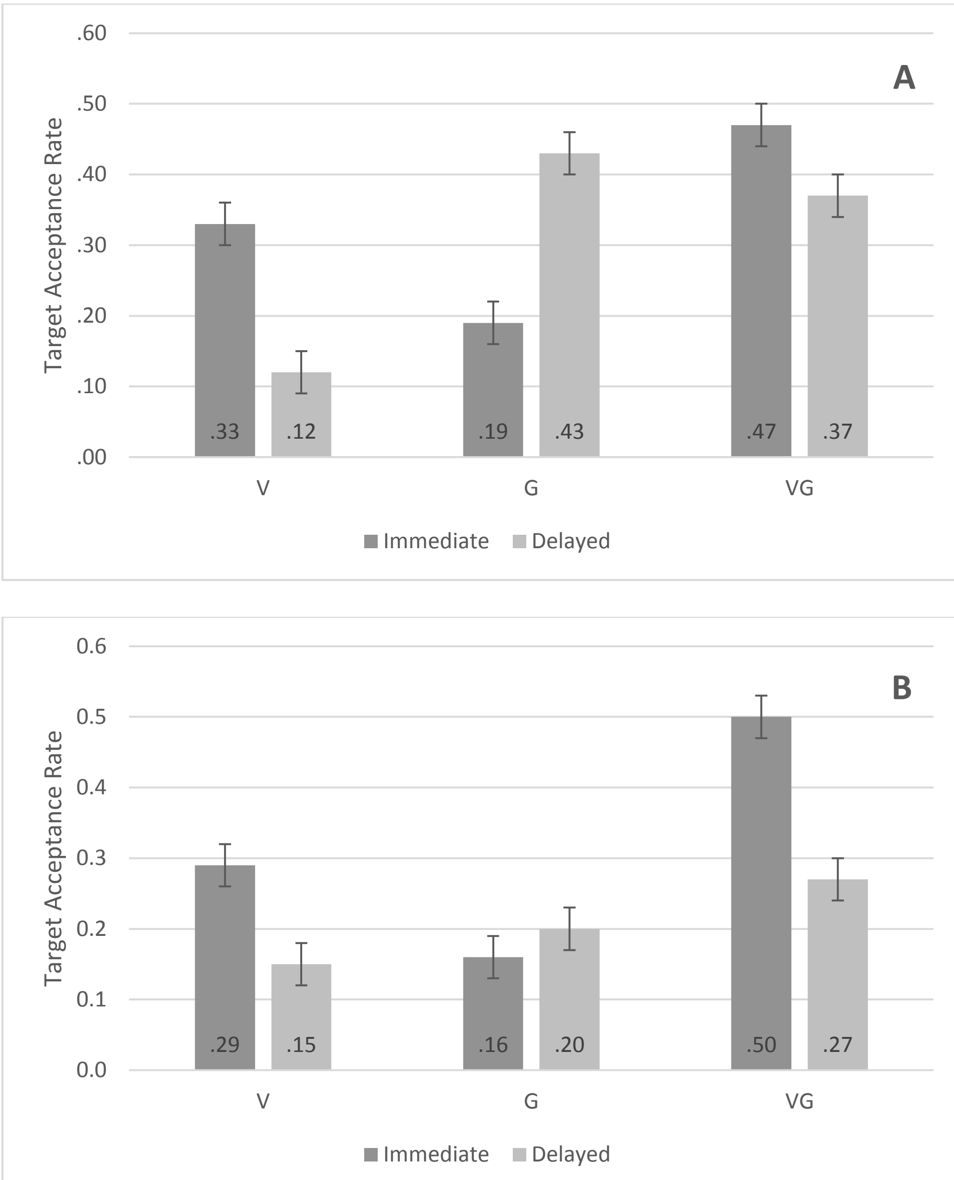


Figure 4.6. Forgetting of targets by question for short lists (panel A) and long lists (panel B) in Experiment 2.

As for critical distractors, the main effect of test and the question x test interaction were reliable, and additionally there was a question x length x test interaction. The main effect of test was similar to the effect on targets because the critical distractor acceptance rate decreased from the immediate (M = .41) to the delayed test (M = .29), $F(1,86) = 13.75$, $p < .05$. The question x test interaction was similar—but not identical—to the effect for targets: CD acceptance in the V and VG conditions declined over the delay while CD acceptance in the G remained stable over the delay, $F(2,172) = 15.43$, $p < .05$. The effect of forgetting and list length on CDs was not the same as the effect on targets: Forgetting did not impact long lists (.41 vs. .36), but reduced CD acceptance for short lists (.40 vs. .23). Finally, the question x length x test interaction, revealed that subjects experienced forgetting in all conditions for short lists, but for long lists they remembered more CDs over time in the G condition (see Figure 4.7), $F(2,172) = 7.84$, $p < .05$.

Summary. As in Experiment 1, negative moods reduced standard false memory compared to positive moods on the immediate test. Unlike Experiment 1, mood did not affect true memory on the delayed test. The new list length manipulation affected false memory, as predicted, but only on the delayed test, and did not impact true memory. Longer lists resulted in more false memories than shorter lists after a delay, suggesting that longer lists strengthened gist memory more than short lists. Forgetting occurred for both true and false memories, but the effect on false memories was driven by short lists.

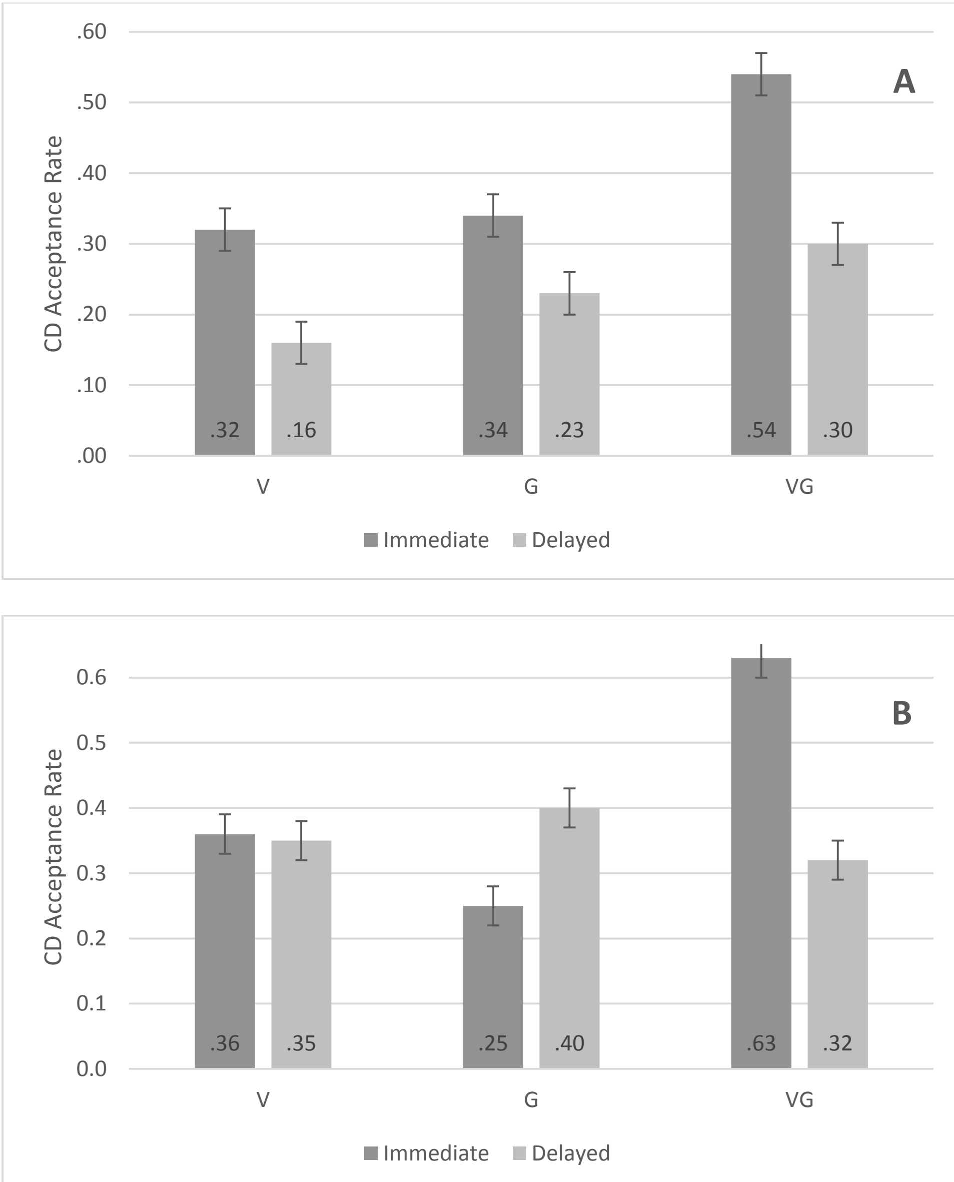


Figure 4.7. Forgetting of CDs by question for short lists (panel A) and long lists (panel B) in Experiment 2.

When all instructional conditions were considered, list length still did not affect target acceptance on the delayed test, but CD acceptance in the G condition was higher for short lists than long lists on the immediate test. On the delayed test, target acceptance was higher for short lists compared to long lists, whereas the opposite was true for CDs. The effect of repeated testing was also conditional on list length and the pattern was identical for targets and CDs:

Testing items twice increased acceptance for short lists—in all instructional conditions—but testing items twice reduced acceptance for long lists in the G condition.

Finally, even though forgetting resulted in lower levels of standard true and false memory over time, target acceptance in the G condition actually increased over time and CD acceptance in the G condition remained stable. Furthermore, forgetting occurred for short lists but not for long lists, for both targets and CDs. The increase in G condition acceptance for targets was driven by short lists, whereas the stability of G condition acceptance for CDs occurred because of increased G acceptance for long lists.

Model Fits

The model's fit was evaluated by computing likelihood ratio tests, as was done in Experiment 1. The G^2 values for all of the conditions were again below 3.84, indicating that the model provided a statistically acceptable estimation of the data.

Parameter Estimates

Negative moods increased recollection rejection and reduced familiarity on the immediate test, consistent with the finding that negative moods reduced false memory. Similarly, longer lists reduced recollection rejection and increased familiarity compared to shorter lists on the delayed test, consistent with the fact that standard false memory was higher for longer lists. Target acceptance was higher for short lists on the delayed test, which was driven by increased identity judgment.

Table 4.3

Experiment 2 Parameter Estimates

	R_T	F_T	P_D	R_D	F_D	β_V	β_G	β_{VG}
Immediate Test								
Short Positive	.35	.19	.27	.26	.44	.42	.42	.29
Short Negative	.34	.19	.52	.26	.05	.24	.45	.24
Short Neutral	.31	.23	.24	.18	.41	.25	.39	.25
Long Positive	.37	.03	.34	.21	.49	.42	.42	.29
Long Negative	.31	.20	.53	.26	.46	.24	.45	.24
Long Neutral	.31	.19	.34	.08	.40	.25	.39	.25
Positive	.36	.11	.30	.22	.46	.42	.42	.29
Negative	.32	.19	.47	.26	.31	.24	.45	.24
Neutral	.31	.21	.31	.11	.40	.25	.39	.25
Short	.33	.20	.33	.21	.34	.28	.43	.28
Long	.33	.16	.39	.17	.44	.27	.43	.27
Mean	.33	.18	.35	.17	.40	.27	.43	.27
Delayed Test								
Short Positive	.23	.11	.19	.18	.30	.41	.39	.39
Short Negative	.19	.20	.24	.18	.19	.36	.31	.36
Short Neutral	.24	.11	.25	.24	.01	.40	.40	.41
Long Positive	.11	.21	.08	.03	.36	.41	.39	.39
Long Negative	.18	.18	.33	.21	.24	.37	.34	.34
Long Neutral	.20	.18	.27	.13	.29	40	.40	.41
Positive	.16	.18	.13	.11	.33	.41	.39	.39
Negative	.17	.20	.27	.19	.22	.36	.31	.36
Neutral	.22	.15	.26	.19	.16	.40	.40	.41
Short	.21	.14	.22	.20	.17	.39	.37	.39
Long	.16	.20	.21	.11	.31	.38	.38	.38
Mean	.19	.17	.22	.16	.24	.38	.38	.38
Combined Across Both Tests								
Positive	.24	.16	.19	.14	.37	.41	.41	.35
Negative	.23	.19	.35	.21	.23	.33	.35	.31
Neutral	.25	.17	.29	.16	.24	.35	.39	.35
Short	.30	.20	.21	.06	.02	.28	.43	.28
Long	.22	.19	.05	.14	.50	.38	.38	.34
Short Positive	.28	.12	.24	.21	.34	.41	.41	.35
Short Negative	.24	.18	.25	.23	.15	.35	.35	.31
Short Neutral	.27	.14	.19	.24	.12	.35	.39	.35
Long Positive	.21	.18	.15	.08	.38	.41	.41	.35

Long Negative	.21	.19	.37	.23	.31	.35	.35	.31
Long Neutral	.24	.19	.30	.11	.33	.35	.39	.35

Table 4.4

Experiment 2 Prior Testing Parameter Estimates

	R_T	F_T	P_D	R_D	F_D	β_V	β_G	β_{VG}
Not Previously Tested								
Positive	.14	.04	.11	.09	.26	.41	.39	.38
Negative	.17	.03	.16	.19	.21	.36	.31	.36
Neutral	.21	.04	.14	.21	.09	.40	.40	.41
Short	.21	.00	.09	.14	.09	.39	.37	.39
Long	.13	.07	.06	.09	.39	.39	.37	.39
Short Positive	.22	.00	.09	.11	.18	.41	.39	.39
Short Negative	.17	.00	.21	.19	.09	.36	.31	.36
Short Neutral	.23	.04	.00	.12	.00	.39	.39	.41
Long Positive	.06	.08	.08	.00	.38	.42	.38	.38
Long Negative	.17	.06	.00	.16	.39	.36	.31	.36
Long Neutral	.18	.04	.09	.13	.38	.40	.40	.41
Mean	.17	.04	.13	.16	.19	.39	.37	.39
Previously Tested								
Positive	.18	.30	.14	.12	.42	.41	.39	.39
Negative	.17	.35	.38	.19	.23	.36	.31	.36
Neutral	.23	.25	.38	.16	.24	.40	.40	.41
Short	.29	.28	.30	.14	.42	.28	.43	.28
Long	.26	.37	.42	.17	.31	.28	.43	.28
Short Positive	.23	.21	.14	.13	.50	.41	.39	.39
Short Negative	.19	.37	.18	.10	.36	.36	.31	.36
Short Neutral	.25	.16	.35	.18	.25	.40	.40	.41
Long Positive	.17	.35	.07	.06	.35	.41	.39	.39
Long Negative	.15	.31	.67	.31	.01	.36	.31	.36
Long Neutral	.24	.31	.41	.16	.23	.40	.40	.41
Mean	.27	.33	.37	.15	.35	.28	.43	.28

Consistent with the finding that repeated testing increased target and CD acceptance in all conditions for short lists, testing items twice increased identity and familiarity parameters, but did not increase recollection rejection (see Table 4.4). On the other hand, testing items twice reduced familiarity for long lists, consistent with the reduction in CD acceptance in the G condition for long lists.

The parameter results were fairly consistent with the ANOVAs for forgetting as well. Estimates of identity judgment declined over time, which is consistent with the reduction in true memory. For short lists, CD acceptance declined in the V and G condition, and recollection rejection and familiarity both declined over time. For long lists, CD acceptance in the G condition remained stable over time, and familiarity remained stable while recollection rejection decreased with time.

Discussion

Experiment 2 showed that negative moods reduce false memory compared to positive moods immediately after list presentation, as in the first study. Experiment 1 also showed that negative moods increased true memory after a delay, an effect that was not replicated here. The true memory result in Experiment 1 may have been simply an artifact of bias correction because target acceptance in the V condition was unusually low. Experiment 2 also provided the additional new result that longer lists increased false memory after a delay by reducing recollection rejection and increasing familiarity. Again, this result is consistent with the established idea that shorter lists favor verbatim processing and longer lists favor gist processing (Sugrue & Hayne, 2006).

Although at first glance it appears that the effect of mood on false memory was the same in the present study as it was in the first study, inspection of the parameters reveals otherwise. Recall that in Experiment 1, negative mood reduced false memory because it increased identity judgment compared to positive moods; none of the other parameters were reliably different as a function of mood. In Experiment 2, however, negative mood increased recollection rejection and reduced familiarity, but did not impact identity judgment. The only difference between Experiment 1 and Experiment 2 was the length of the study lists and indeed, further unpacking of the parameter estimates provides some evidence that the different processes now causing mood effects are related to the new list lengths. In that connection, the reduction in familiarity for negative moods was entirely driven by short lists. Conversely, the increase in recollection rejection was primarily driven by long lists more than short lists. On short lists negative moods increased recollection rejection compared to neutral moods, but they did not increase recollection rejection compared to positive moods. On long lists, however, negative moods increased recollection rejection compared to positive moods and especially compared to neutral moods.

This processing difference between the first and second studies—even though the effect on false memory was the same—suggests that subjects using a verbatim processing style as a result of negative mood will be affected by whether the study items further encourage verbatim processing or whether they encourage gist processing instead. When encouraged to use verbatim processing on short lists, subjects in negative moods will have extremely low levels of familiarity because verbatim traces are particularly strong. On the other hand, subjects in negative moods who are encouraged to use gist processing by long lists will be better able to reject CDs than their counterparts using gist processing with less ability to reject CDs.

Across list lengths, the effects of prior testing and forgetting were the same as in Experiment 1, and the associated parameter differences were similar, especially for prior testing: Repeated testing increased target and CD acceptance by increasing identity judgment, phantom recollection, and familiarity, while not impacting recollection rejection. Forgetting reduced identity judgment, as in Experiment 2, but also reduced recollection rejection and familiarity.

However, repeated testing impacted memory for short and long lists differently, because it increased G condition acceptance for short lists, but it reduced G condition acceptance for long lists. The different pattern for short and long lists was primarily driven by the fact that familiarity was much higher for untested items from long lists compared to short lists, such that short lists had more room for an increase and long lists had more room for a decrease in familiarity as a result of repeated testing. This effect remains consistent with the relative contributions of verbatim and gist processing to short and long lists, respectively, as well as the idea that repeated testing is a verbatim effect. For long lists, gist traces will be stronger initially due to the many associated targets presented. If verbatim traces are strengthened by repeated testing it may result in lower G condition acceptance. For short lists, on the other hand, gist traces will be much weaker initially and may be strengthened by repeated testing.

Similarly, although the overall forgetting effect was the same as in Experiment 1, forgetting impacted short and long lists differently. Forgetting of both targets and CDs occurred for short lists but not long lists, which was driven by declines in identity judgment on the target side and declines in familiarity on the CD side. It appears that the stronger gist activation created by longer lists was more robust to forgetting. Furthermore, target acceptance in the G condition increased over time and CD acceptance in the G condition remained stable. If familiarity is strong for CDs right away, these strong gist traces would remain intact over two days. In the case of targets, it is possible that the decay of verbatim memory traces with time allowed for increased retrieval of gist traces. This would lead to increased acceptance in the G condition if verbatim traces were unavailable and subjects could not remember that a target had been presented, but did remember the theme of the list.

The results of Experiment 2 are generally consistent with the hypotheses that negative moods promote verbatim processing and that the additional manipulations are generally verbatim manipulations as well. Mood effects have been isolated to the immediate test and it is possible that a stronger mood manipulation could create more robust mood effects that persist over the two-day delay period. Alternatively, the addition of increased arousal may affect the pattern of mood effects on the immediate test. If increased arousal increases false memory to the extent that negative moods no longer reduce false memory compared to neutral moods, it would indicate a valence-arousal interaction that is consistent with predictions of FTT. If increased arousal reduces false memory for negative moods but increases false memory for positive moods, that would be consistent with predictions of AAI. The third experiment addressed these possibilities via a mood manipulation that created higher levels of arousal.

CHAPTER 5

EXPERIMENT 3: EFFECT OF HIGH-AROUSAL MOOD ON TRUE AND FALSE MEMORY FOR NEUTRAL WORDS

In Studies 1 and 2, mood effects were limited to the immediate test, and positive moods increased false memory while negative moods decreased false memory, relative to neutral moods. Experiment 3 tested the hypothesis that mood effects on false memory could be strengthened or altered by increasing arousal from low levels to moderate to high levels. It also addressed the prediction that increasing arousal would affect the processes contributing to true and false memory as a function of valence.

Method

Participants

98 subjects participated in exchange for course credit (71 women). Their ages ranged from 18-27 (M_{age} = 20.41, SD = 1.60). 34 subjects were in the negative group, 37 were in the neutral group, and 27 were in the positive group after excluding those whose mood was inconsistent with the manipulation.

Materials and Procedure

The mood induction materials were identical to those used in Pilot 3 (note that in Experiment 1 and Experiment 2 the mood induction materials from Pilot 2 were used). Aside from the videos used in the mood induction, the materials and procedure were identical to those used in Experiment 2. That is, the same short and long word lists were used, but the high arousal positive and negative videos were used in place of the low arousal videos. The neutral video remained the same.

Results

Mood Manipulation Check

ANOVAs were conducted to determine whether the mood manipulation resulted in a) reliable valence differences between all three groups, b) higher arousal in the negative and positive groups compared to the neutral group, and c) equivalent arousal in the positive and negative groups. The results confirmed that the manipulation was effective in satisfying all three conditions. Post-manipulation valence ratings were highest for the positive group (M = 7.33, SD = .82), followed by the neutral group (M = 5.22, SD = .51), which was followed by the negative group (M = 2.44, SD = .71), $F(2,97) = 33.65$, $p < .05$. Note that the valence levels for the positive and neutral group are equivalent to those found in Experiment 1 and Experiment 2, but the negative group in the present study experienced a slightly lowered valence level than those in the previous two studies. This difference is likely a result of the elevated arousal levels, as valence and arousal are sometimes correlated at more extreme valences (see Mattek, Wolford, & Whalen, 2017). In other words, when arousal is increased negative moods may become more strongly negative.

Post-manipulation arousal ratings were higher for the negative group (M = 5.71, SD = 1.03) and the positive group (M = 6.04, SD = 1.67) than the neutral group (M = 4.30, SD = .73), $F(2,97) = 19.21$, $p < .05$, but arousal did not differ between the two emotional groups. Note that these levels of arousal are higher than those found for the positive and negative valence groups in Experiment 1 and Experiment 2.

Bias correction

The effects of conjoint recognition question and mood on unrelated distractor acceptance were identified. Two parallel 3 (question: V, G, VG) x 3 (mood: positive, negative, neutral) analyses of variance (ANOVAs) were performed on the acceptance rates for unrelated distractors, separately for the immediate and the delayed tests. There was a reliable main effect of question on the immediate test but not the delayed test, and no effects of valence on either test. On the immediate test, unrelated distractor acceptance was higher in the VG condition than the V and G conditions, which did not differ from each other, $F(2,302) = 10.08$, $p < .05$ (see Table 3.3 for descriptive statistics). Because the unrelated distractor acceptance was affected by question, and because the effect of question was not consistent over the delay, the data were bias-corrected using the unrelated distractor acceptance rates separated by question and test. All of the following analyses were conducted on the bias-corrected values.

Analysis of Standard True and False Memory

Neither mood nor length affected true memory on the immediate test. False memory, however, was affected by the interaction between list length and valence on the immediate test, $F(2,95) = 6.28$, $p < .05$. There were no reliable mood effects on false memory for long lists, but for short lists positive moods (.62) increased false memory compared to negative moods (.47), which increased false memory compared to neutral moods (.29). There was no main effect of valence on false memory, unlike in Experiment 1 and 2.

There were no main effects of valence on true or false memory on the delayed test. However, forgetting occurred for true memory, as it declined from .32 on the immediate test to .17 on the delayed test, $F(1.95) = 30.99$, $p < .05$. Forgetting occurred for false memory as well, as it declined from .42 to .25, $F(1,95) = 27.88$, $p < .05$. This effect was driven by short lists because false memory for short lists declined over time (.46 vs. .19) but false memory for long lists remained stable over time (.38 vs. .32), $F(1,95) = 14.39$, $p < .05$. There was also a marginal main effect of valence in the forgetting analysis, because subjects in the positive group (.41) had more false memories than subjects in the neutral group (.27), $F(1,95) = 3.17$, $p = .05$.

Main ANOVA Results

Four separate 3 (question: V, G, VG) x 3 (mood: negative, neutral, positive) x 2 (list length: short, long) ANOVAs were conducted on target and CD bias-corrected acceptance rates, separately for the immediate and delayed memory tests, with question and length as within-subjects factors and mood as the between-subjects factor.

Immediate test. Turning to targets first, the main effect of question revealed that target acceptance was highest in the VG condition ($M = .45$), followed by the V condition ($M = .32$), then the G condition ($M = .19$); all mean differences were reliable, $F(2,302) = 29.27$, $p < .05$. There were no mood effects.

Moving on to CDs, acceptance was highest in the VG condition ($M = .55$), followed by the V condition ($M = .42$), and finally the G condition ($M = .27$); all mean differences were reliable, $F(2,190) = 24.25$, $p < .05$. Furthermore there was a length x valence interaction, indicating neutral moods reduced CD acceptance compared to emotional moods for short lists but not for long lists (see Figure 5.1), $F(2,95) = 5.21$, $p < .05$.

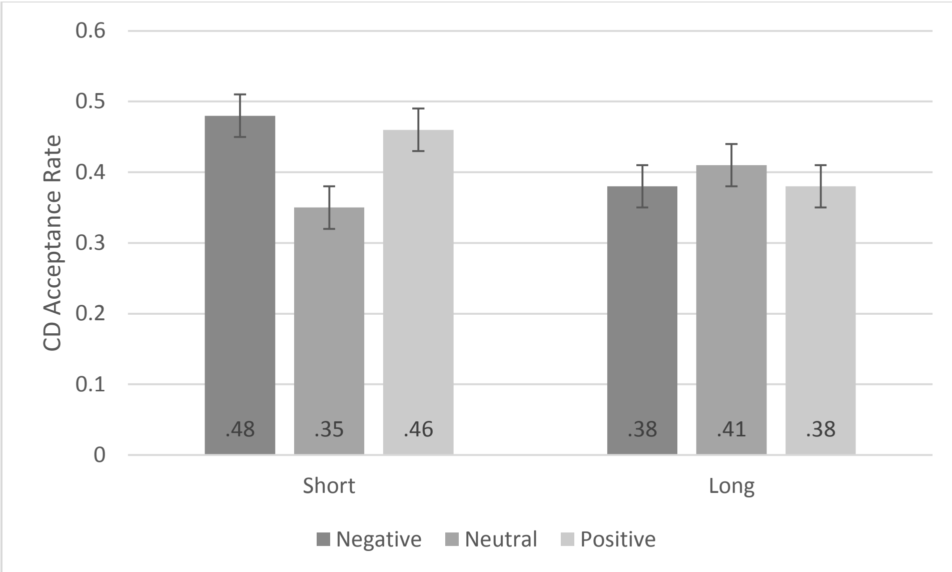


Figure 5.1. CD acceptance by valence and list length on the immediate test in Experiment 3.

Finally, the three-way interaction was marginally significant, showing that positive moods increased CD acceptance in the V condition compared to negative and neutral moods, but only for short lists (see Figure 5.2), $F(4,190) = 2.48$, $p = .05$.

Delayed test. For targets there was a main effect of question, $F(2,302) = 12.65$, $p < .05$, revealing that the effect of question was such that target acceptance was greater in the VG condition ($M = .33$) than the V ($M = .24$) and G ($M = .19$) conditions, which did not differ. Target acceptance was higher for short lists (.30) than long lists (.26) as in Experiment 2, $F(1.95) = 11.44$, $p < .05$. Finally, shorter lists increased target acceptance in the G condition but not in other conditions, as in Experiment 2, $F(2,190) = 12.12$, $p < .05$.

For CDs there was a reliable main effect of question, $F(2,188) = 20.87$, $p < .05$, such that CD acceptance was greater in the VG condition ($M = .42$) than the V ($M = .31$) and G conditions ($M = .27$). No other effects were reliable.

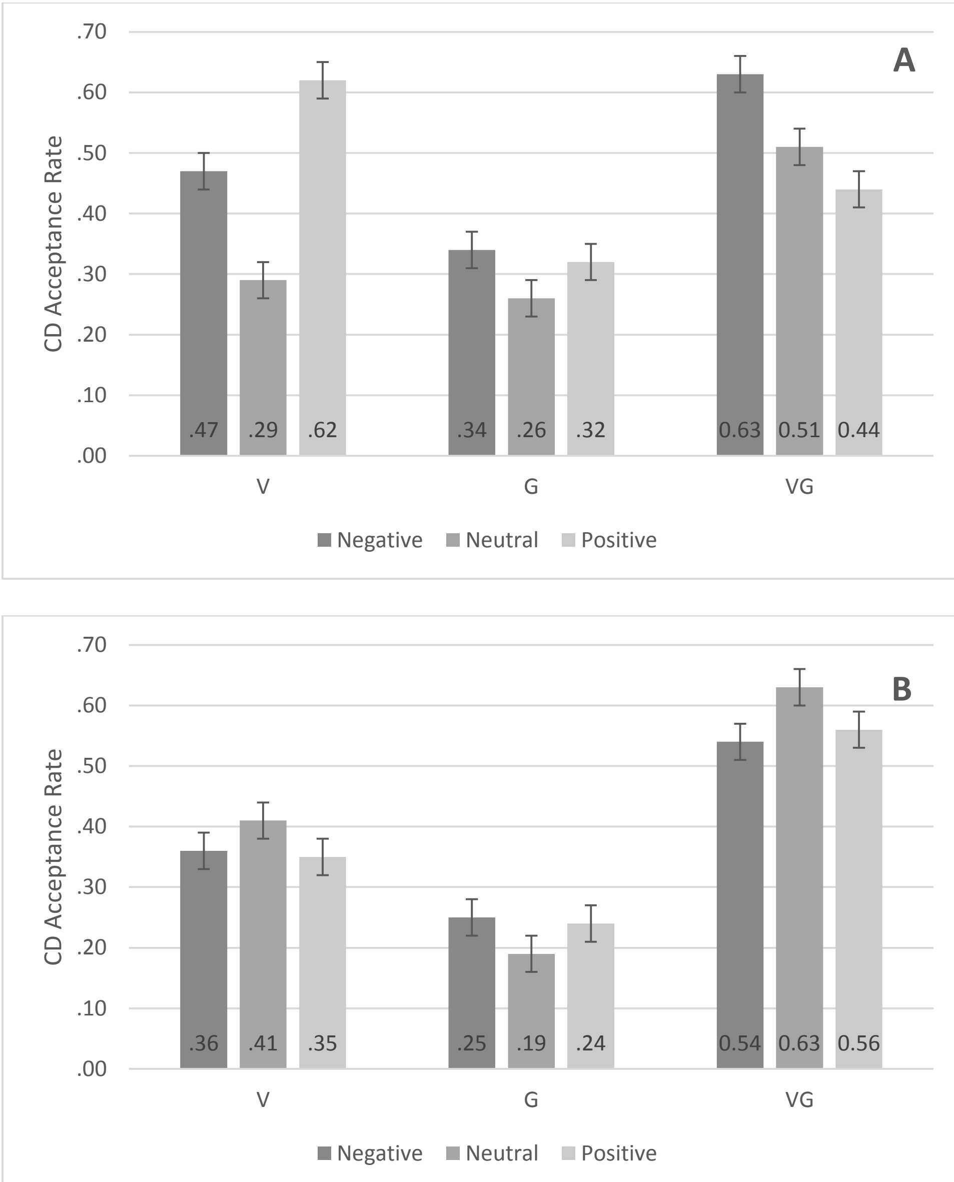


Figure 5.2. CD acceptance by valence and question for short lists (panel A) and long lists (panel B) on the immediate test in Experiment 3.

Prior testing. Another set of ANOVAs were conducted in order to compare acceptance rates of targets and critical distractors on the delayed test that had not been tested on the immediate test to targets and critical distractors on the delayed test that had been tested on the immediate test. That is, this comparison was an analysis of the effect of prior testing, and applies only to items on the delayed test. The design was thus two parallel 3 (question: V, G, VG) x 3 (mood: negative, neutral, positive) x 2 (length: short, long) x 2 (times tested: once, twice) ANOVAs conducted on delayed test target and CD acceptance rates.

Turning first to targets, there was a reliable effect of prior testing, because targets that were tested twice ($M = .29$) were more likely to be accepted than targets that were only tested once ($M = .22$), $F(1,151) = 41.99$, $p < .05$. Furthermore, there were several reliable interactions. The question x prior testing interaction revealed that testing targets twice increased target acceptance in the V condition but in not the G condition, where the reverse was true, $F(2,302) = 7.58$, $p < .05$. Prior testing increased target acceptance for long lists (.22 vs. .29) but not short lists (.31 vs. .30), $F(1,95) = 14.98$, $p < .05$. Note that in Experiment 2 prior testing increased target acceptance for long lists and reduced it for short lists. The three-way interaction, shown in Figure 5.3, indicated that testing items twice increased target acceptance in the V condition for short lists only, but testing items twice *reduced* target acceptance in the G condition for short lists only, $F(2,302) = 8.75$, $p < .05$. Testing items twice increased target acceptance in the VG condition for long lists only. Note that this pattern differs from the one found in Experiment 2 (see Figure 4.3), where testing items twice increased target acceptance in all conditions for short lists, but reduced target acceptance in the G condition for long lists.

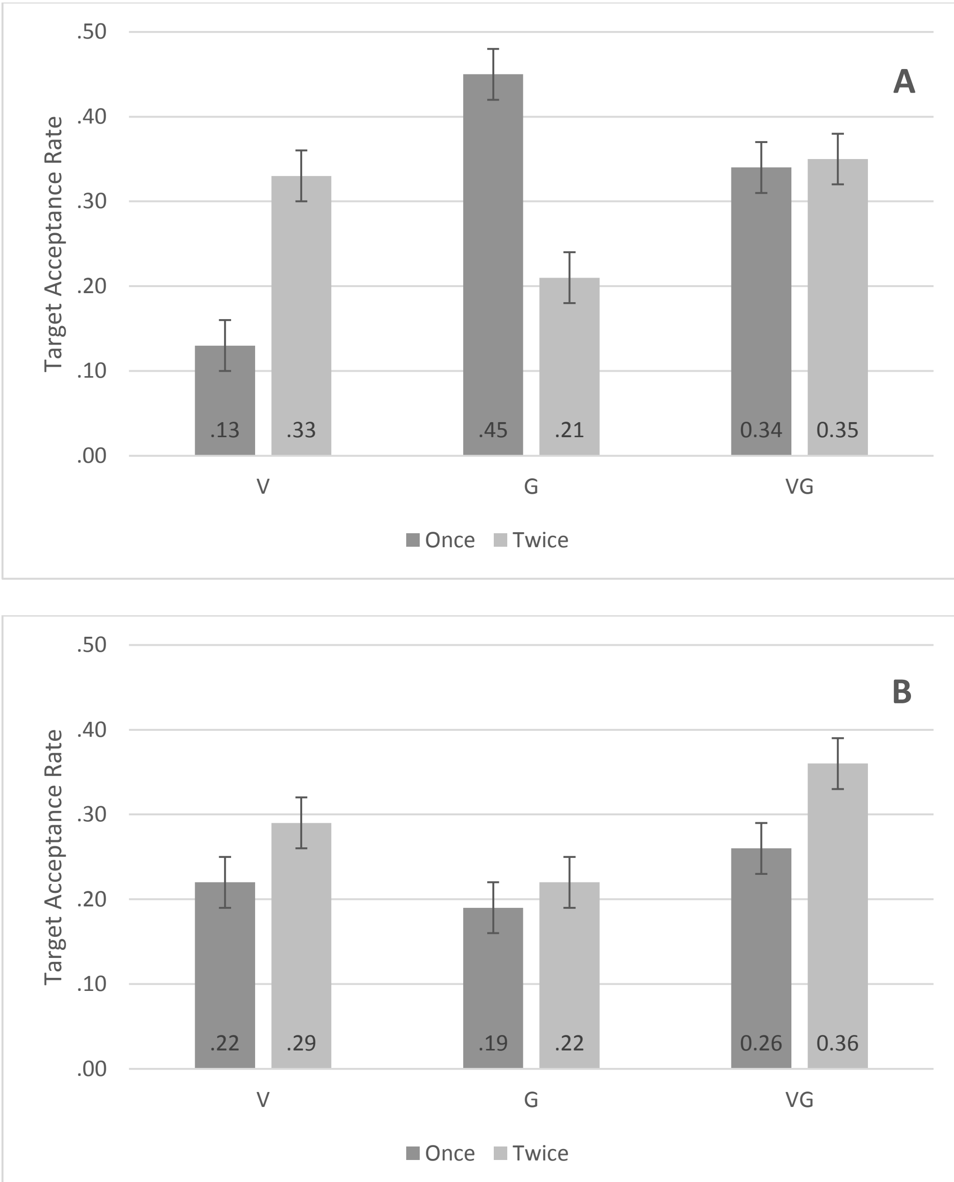


Figure 5.3. Target acceptance rate by prior testing and question for short lists (panel A) and long lists (panel B) on the delayed test (Experiment 3).

Turning now to CDs, the main effect of prior testing was reliable because CDs were more likely to be accepted when tested twice (M = .37) rather than once (M = .30), $F(1,188) = 9.06$, $p < .05$. Furthermore, there were several reliable interactions. First, testing items twice increased CD acceptance in the V and VG conditions, but not the G condition, $F(2,188) = 5.45$, $p < .05$. In terms of length, testing items twice increased CD acceptance for short lists and did not affect CD acceptance for long lists, $F(1,94) = 26.19$, $p < .05$. Prior testing also increased CD acceptance for negative and neutral moods, but not positive moods, $F(2,94) = 2.93$, $p = .06$ (see Figure 5.4). Note that this effect was not present in Experiment 2, when arousal was low.

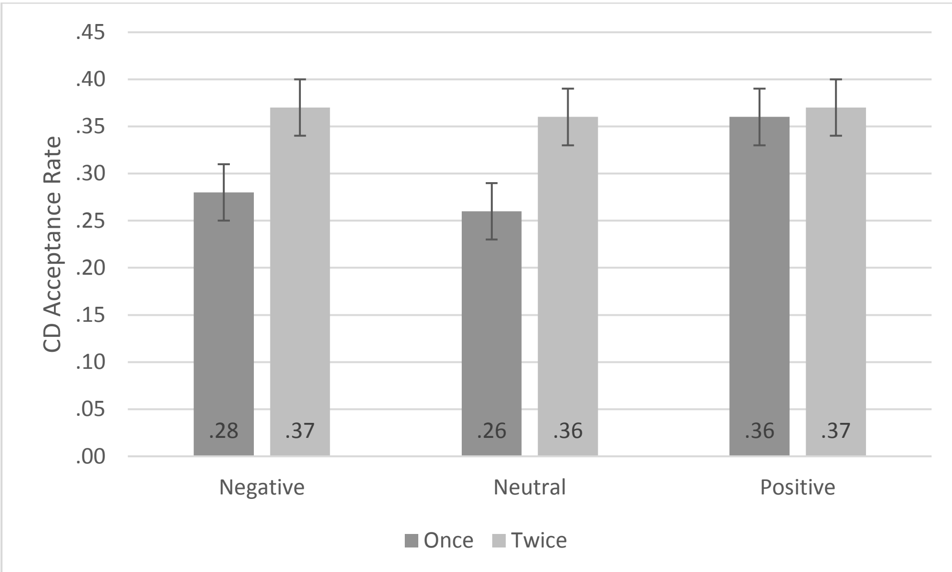


Figure 5.4. CD acceptance rate by mood and prior testing (Experiment 3).

Finally, the three-way interaction between length, question, and prior testing, shown in Figure 5.5, revealed that for short lists testing items twice increased CD acceptance in the V condition, but for long lists testing items twice reduced CD acceptance in the G condition, $F(2,188) = 8.02$, $p < .05$. Note that this interaction is similar to the parallel interaction in Experiment 2 (see Figure 4.4). The main difference is that here, the effect of prior testing was not reliable in the G and VG conditions for short lists, but the increase was reliable in Experiment 2.

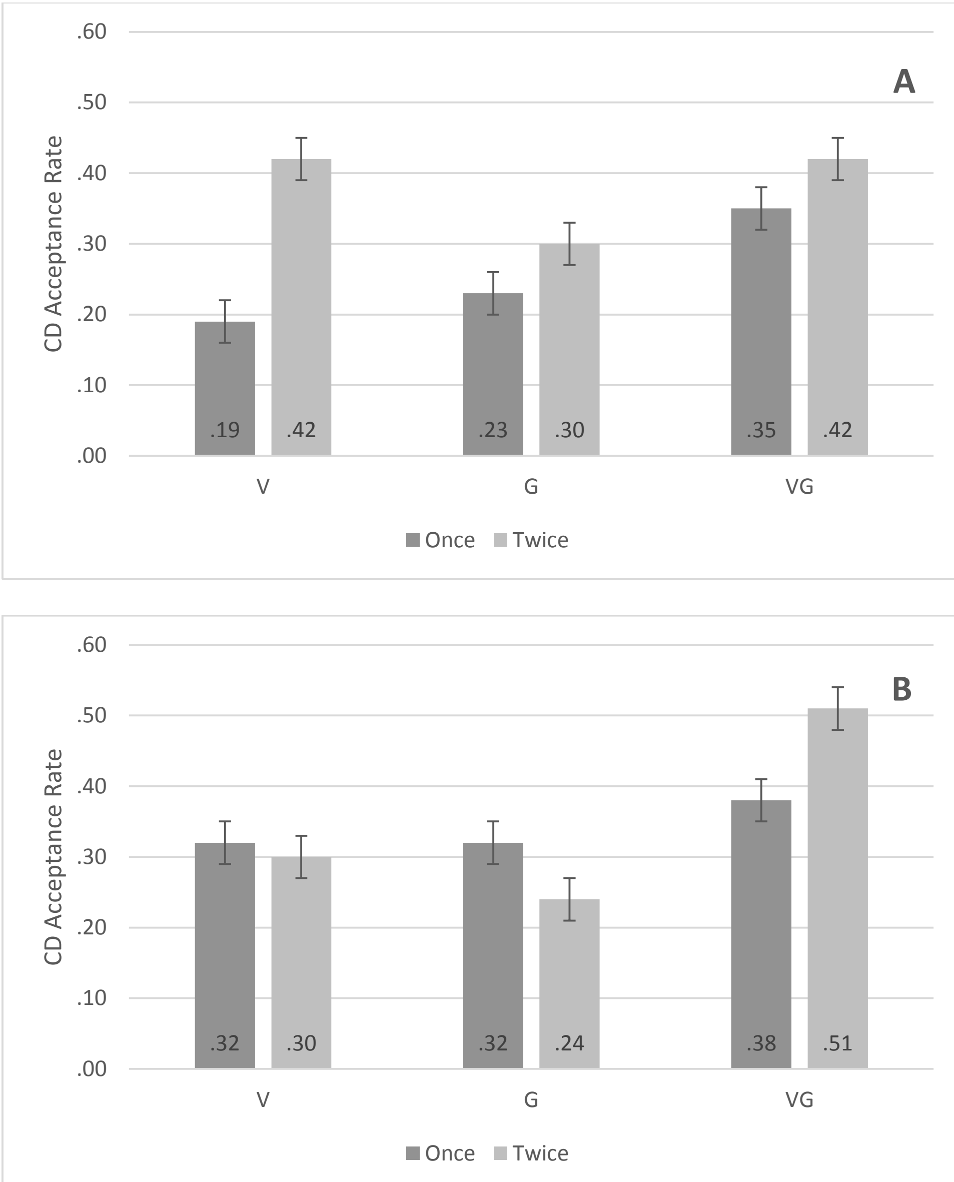


Figure 5.5. CD acceptance rate by question and prior testing for short lists (panel A) and long lists (panel B) on the delayed test (Experiment 3).

Forgetting. In order to measure subjects' forgetting over the two-day delay period, additional ANOVAs were computed with test (immediate vs. delayed) as a factor, excluding items on the delayed test that were being tested for a second time. That is, a comparison was made between items that were tested immediately after the study phase and items tested for the first time two days after the study phase. The design was thus two parallel 3 (question: V, G, VG) x 3 (mood: negative, neutral, positive) x 2 (length: short, long) x 2 (test: immediate, delayed) ANOVAs conducted on target and CD acceptance rates.

The target acceptance rate was higher on the immediate test ($M = .32$) than on the delayed test ($M = .26$), $F(1,95) = 29.98$, $p < .05$. Note that this is the same pattern observed in Experiment 1 and Experiment 2. Furthermore, the test x question interaction was consistent with the previous experiments: In the V and VG conditions, target acceptance was higher on the immediate test than on the delayed test but in the G condition target acceptance increased over the delay, $F(2,190) = 14.17$, $p < .05$ (see Figure 5.6, panel A). Target acceptance furthermore declined over the delay for long lists (.33 vs. .22) but not short lists (.31 vs. .31), as in Experiment 2, $F(1,95) = 20.84$, $p < .05$.

Finally, the three-way interaction, shown in Figure 5.7, revealed that forgetting reduced target acceptance in the V and VG conditions for both list lengths, but in the G condition target acceptance increased over time for short lists and remained stable for long lists, similar to the effect in Experiment 2, $F(2,190) = 15.44$, $p < .05$.

As for critical distractors, the main effect of test, the question x test interaction, the length x test interaction, and the three-way interaction were all reliable. The main effect of test was similar to the effect on targets because the critical distractor acceptance rate decreased from the immediate ($M = .41$) to the delayed test ($M = .30$), $F(1,150) = 28.45$, $p < .05$. The question x test interaction was similar to the effect for targets: CD acceptance in the V and VG conditions declined over the delay while CD acceptance in the G condition remained stable over the delay (see Figure 5.6, panel B), $F(2,300) = 8.91$, $p < .05$. The length x test interaction occurred because CD acceptance was reduced over the delay for short lists (.43 vs. .26) but not for long lists (.40 vs. .34), $F(1,94) = 11.94$, $p < .05$.

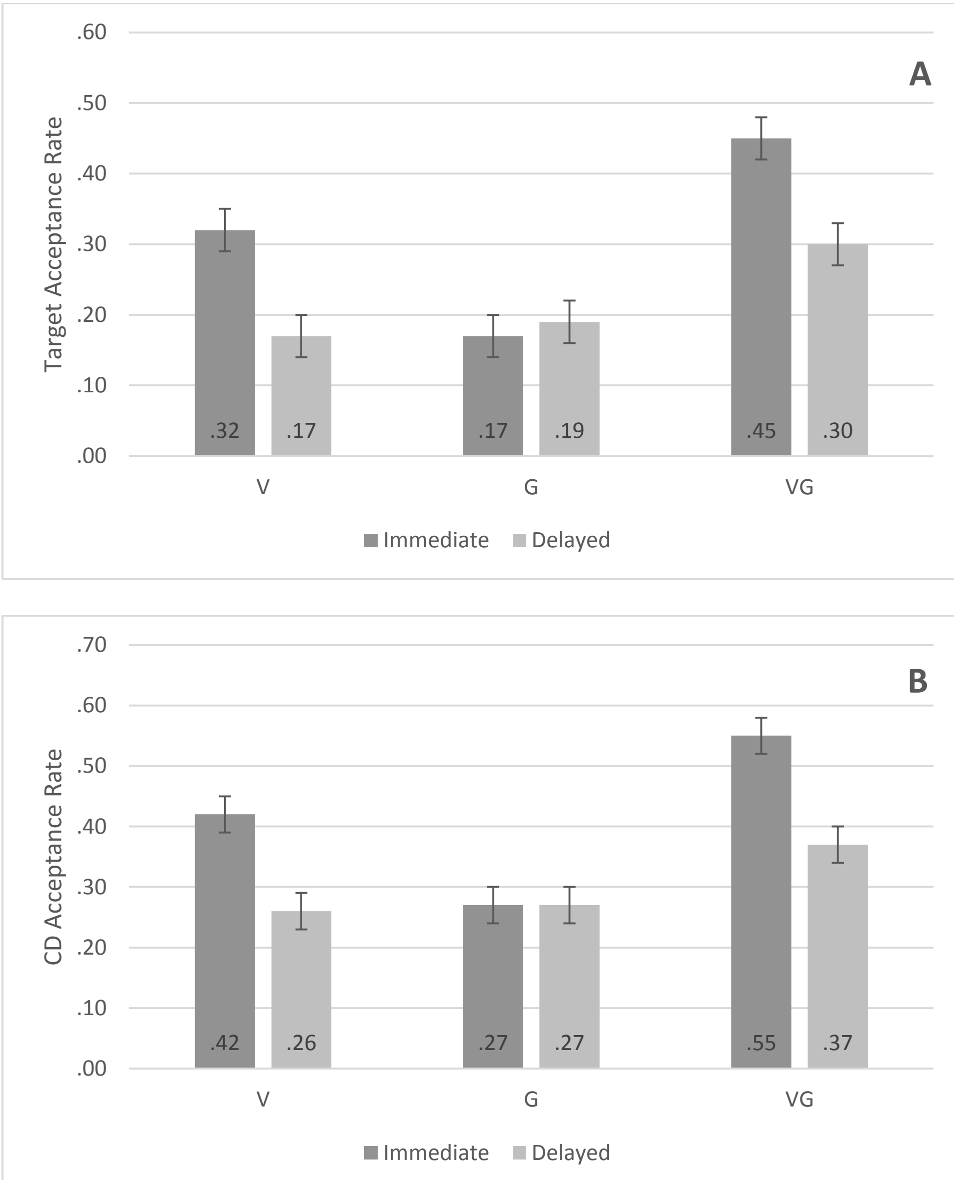
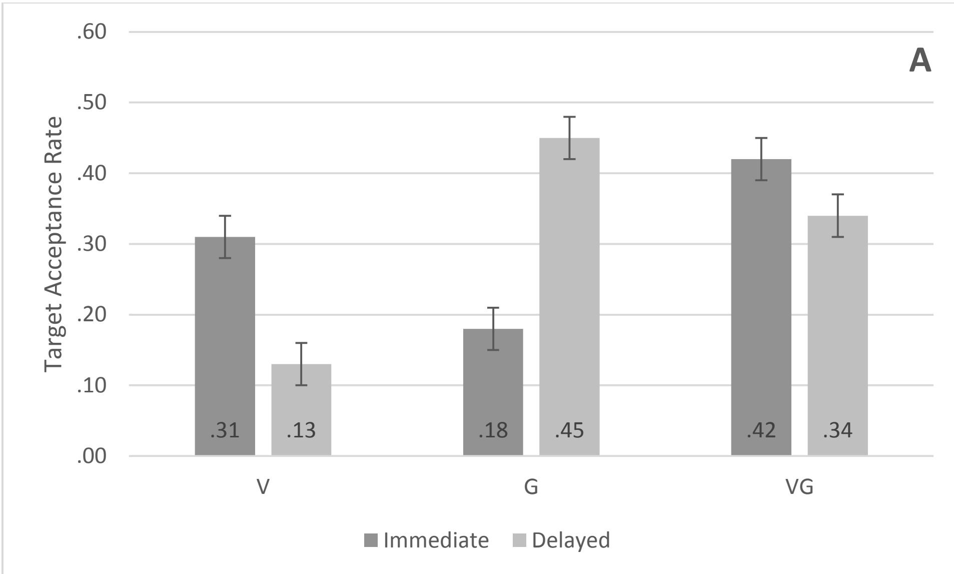


Figure 5.6. Forgetting of targets (panel A) and CDs (panel B) by question in Experiment 3.



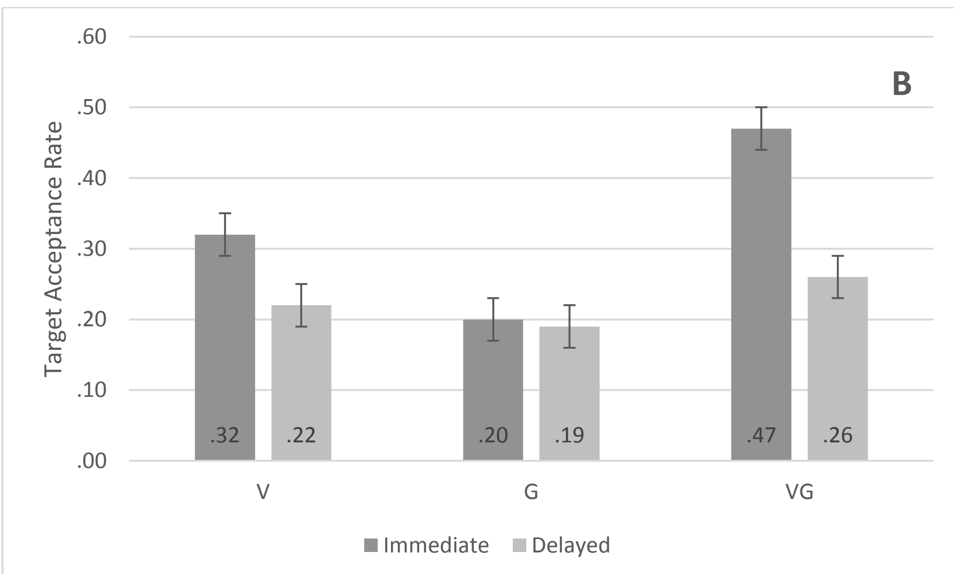


Figure 5.7. Forgetting of targets by question for short lists (panel A) and long lists (panel B) in Experiment 3.

Furthermore, the three-way question x length x test interaction was also reliable, $F(2,188) = 4.47$, $p < .05$ (see Figure 5.8). CD acceptance declined in all conditions for short lists, but declined only in the VG condition for long lists. Note that this pattern is quite similar to the parallel interaction in Experiment 2 (see Figure 4.7).

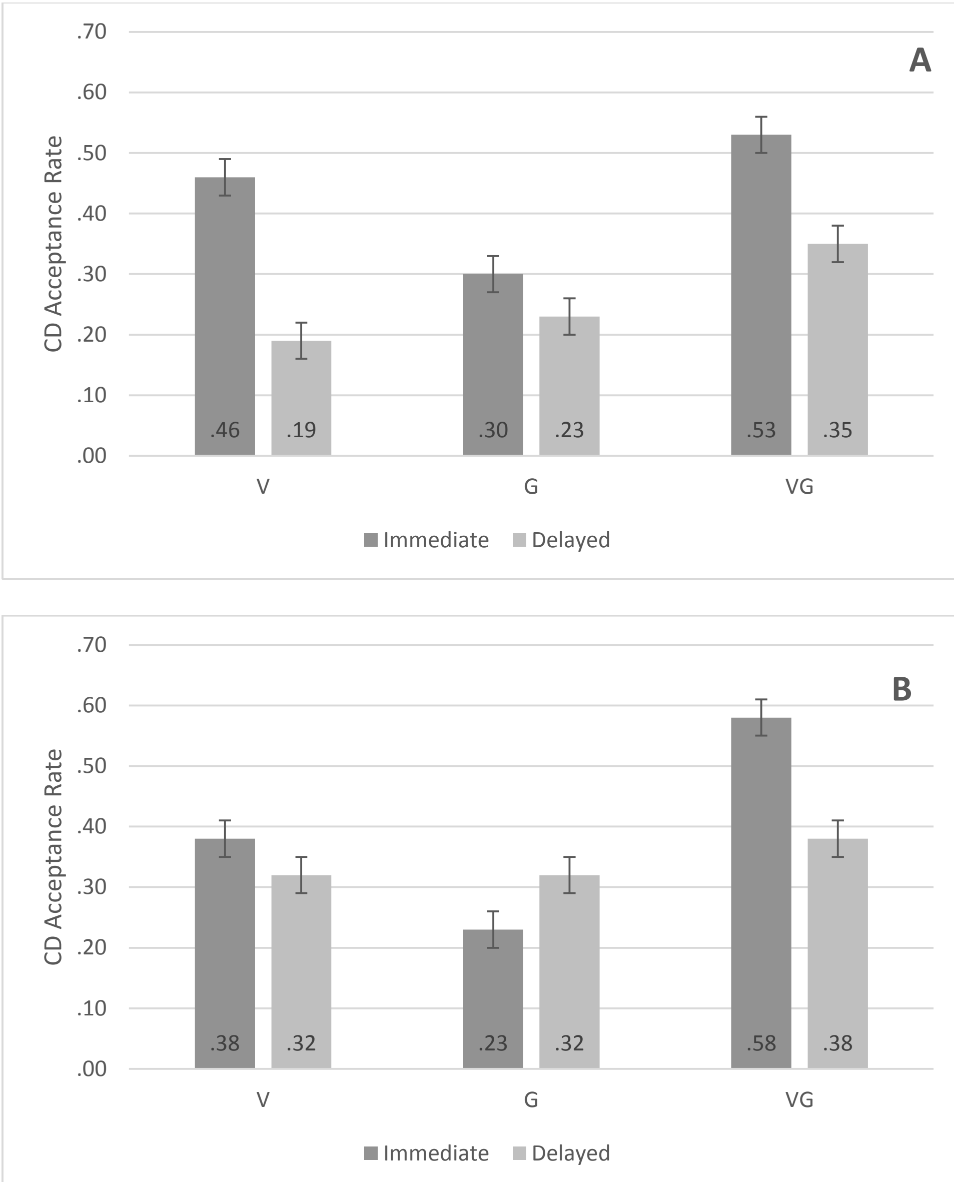


Figure 5.8. Forgetting of CDs by question for short lists (panel A) and long lists (panel B) in Experiment 3.

Summary. Valence did not affect standard true memory in Experiment 3, consistent with prior research, but there were false memory effects on the immediate test that differed from the previous two studies. Valence did not affect false memory for long lists, but for short lists subjects in positive moods experienced more false memory than subjects in negative moods, who experienced more false memory than subjects in neutral moods on the immediate test. Mood valence did not affect CD acceptance in the G and VG conditions, suggesting that mood effects on false memory are related to verbatim retrieval, particularly when arousal is high.

On the delayed test, target acceptance was higher for short lists than long lists, especially in the G condition, as was the case in Experiment 2. Prior testing increased target acceptance in the V condition for short lists but reduced target acceptance in the G condition for short lists. In Experiment 2, on the other hand, prior testing increased target acceptance in the G condition for short lists but reduced target acceptance in the G condition for long lists. In other words, the effect of repeated testing on G condition target acceptance varied for both list lengths, presumably because of differences in arousal.

As for CDs on the delayed test, there was no length effect such as the one found in Experiment 2. Prior testing, however, had effects on CDs that were similar to Experiment 2, with the important addition of a mood effect in Experiment 3: Prior testing increased CD acceptance for negative and neutral moods, but not positive moods.

The pattern of forgetting of standard true and false memory was identical to the pattern in Experiment 2, although here there was an additional valence effect because false memory was higher for subjects in positive moods than neutral moods. Forgetting in all conditions was also similar to Experiment 2, for both targets and CDs. There was again forgetting of CDs in all conditions for short lists, but here the only decline for long lists occurred in the VG condition. In Experiment 2, CD acceptance also increased in the G condition for long lists; here the pattern of means was the same but the difference was not reliable.

Model Fits

The model's fit was evaluated by computing likelihood ratio tests in the same manner it was done for the previous two studies. As in Experiment 1 and Experiment 2, the G^2 values for all of the conditions were below 3.84, indicating that the model provided a statistically acceptable estimation of the data.

Parameter Estimates

Inspection of the parameter estimates in Table 5.1 reveals some potential explanations for the effect of mood on false memory for short lists on the immediate test. Recollection rejection was lower for positive moods compared to neutral moods and both phantom recollection and familiarity were highest for negative moods. These results suggest that increased false memory in positive moods was driven by weaker verbatim traces whereas increased false memory in negative moods was driven by stronger gist traces. False memory was higher for positive moods than negative moods, suggesting that the verbatim effect was stronger than the gist effect.

Prior testing increased target acceptance in the V condition, and the modeling analysis showed that prior testing increased estimates of identity judgment and familiarity (see Table 5.2). As for CDs, prior testing increased acceptance for negative and neutral moods but not positive moods, an effect that appears to be driven by both phantom recollection and familiarity. Furthermore, prior testing also reduced recollection rejection for negative moods. Thus, increased CD acceptance due to prior testing in negative moods was driven by both verbatim and gist retrieval, whereas it was driven only by gist retrieval in neutral moods.

Finally, forgetting reduced both verbatim and gist CD acceptance for short lists, because familiarity declined over the delay and recollection rejection increased. In Experiment 2, the general forgetting effect was similar, but both recollection rejection and familiarity declined over time.

Table 5.1

Experiment 3 Parameter Estimates

	R_T	F_T	P_D	R_D	F_D	β_V	β_G	β_{VG}
				Immediate Test				
Short Positive	.28	.28	.27	.09	.36	.24	.41	.32
Short Negative	.26	.28	.40	.13	.52	.39	.39	.22
Short Neutral	.22	.20	.25	.18	.32	.31	.35	.27
Long Positive	.31	.21	.41	.17	.41	.24	.37	.37
Long Negative	.32	.23	.20	.11	.45	.38	.29	.22
Long Neutral	.17	.37	.35	.08	.41	.33	.33	.27
Positive	.27	.28	.32	.10	.37	.23	.41	.35
Negative	.27	.26	.28	.12	.49	.38	.41	.22
Neutral	.20	.29	.30	.10	.38	.33	.33	.27
Short	.26	.18	.32	.17	.32	.41	.41	.28
Long	.29	.20	.37	.16	.38	.41	.41	.28
Mean	.27	.26	.31	.11	.40	.29	.39	.29
				Delayed Test				
Short Positive	.14	.26	.32	.22	.14	.40	.34	.40
Short Negative	.18	.22	.27	.18	.25	.41	.38	.38
Short Neutral	.18	.03	.37	.30	.05	.39	.39	.31
Long Positive	.13	.23	.30	.18	.36	.40	.34	.40
Long Negative	.20	.23	.32	.14	.24	.41	.38	.38
Long Neutral	.12	.17	.22	.11	.24	.39	.39	.31
Positive	.14	.24	.31	.20	.26	.40	.34	.40
Negative	.19	.22	.30	.16	.25	.41	.38	.38
Neutral	.15	.11	.27	.20	.17	.39	.39	.31
Short	.17	.18	.33	.23	.14	.38	.38	.36
Long	.15	.21	.27	.13	.29	.38	.38	.36
Tested	.19	.31	.36	.17	.35	.38	.38	.36
Untested	.12	.08	.22	.20	.11	.39	.37	.37
Mean	.17	.22	.29	.18	.23	.39	.37	.37
				Combined Across Both Tests				
Positive	.19	.25	.32	.17	.30	.36	.36	.37
Negative	.23	.24	.31	.15	.33	.39	.39	.33
Neutral	.16	.17	.28	.16	.25	.37	.37	.30
Short	.21	.17	.34	.22	.19	.39	.39	.34
Long	.14	.20	.23	.09	.30	.39	.39	.44

Table 5.2

Experiment 3 Prior Testing Parameter Estimates

	R_T	F_T	P_D	R_D	F_D	β_V	β_G	β_{VG}
				Not Previously Tested				
Positive	.00	.14	.24	.20	.21	.40	.34	.40
Negative	.17	.08	.20	.19	.15	.40	.37	.40
Neutral	.09	.02	.24	.22	.02	.39	.39	.31
Short	.16	.03	.23	.18	.01	.38	.38	.36
Long	.11	.10	.17	.11	.29	.38	.38	.36
Short Positive	.13	.13	.19	.18	.10	.40	.34	.40
Short Negative	.21	.01	.10	.15	.13	.41	.38	.38
Short Neutral	.11	.00	.24	.21	.00	.34	.41	.34
Long Positive	.12	.11	.15	.13	.44	.40	.34	.40
Long Negative	.16	.09	.29	.20	.17	.41	.38	.38
Long Neutral	.04	.10	.18	.12	.18	.39	.39	.31
Mean	.12	.08	.22	.20	.11	.39	.37	.37
				Previously Tested				
Positive	.15	.35	.28	.20	.40	.40	.34	.40
Negative	.19	.36	.36	.11	.37	.40	.37	.40
Neutral	.22	.21	.60	.18	.18	.39	.39	.31
Short	.17	.31	.28	.19	.43	.38	.38	.36
Long	.20	.31	.41	.20	.31	.38	.38	.36
Short Positive	.15	.36	.00	.19	.58	.40	.34	.40
Short Negative	.11	.37	.30	.25	.49	.41	.38	.38
Short Neutral	.25	.16	.71	.14	.00	.39	.39	.31
Long Positive	.16	.33	.46	.25	.32	.40	.34	.40
Long Negative	.25	.36	.36	.12	.28	.41	.38	.38
Long Neutral	.21	.25	.32	.14	.45	.39	.39	.31
Mean	.18	.31	.36	.17	.35	.38	.38	.36

Discussion

Mood affected false memory for short lists but not long lists on the immediate test. False memory was highest for subjects in positive moods, followed by negative moods, followed by neutral moods. In the first two studies, false memory was higher for positive moods than negative moods, with neutral moods falling in between. Therefore, increased levels of arousal caused an increase in false memory, lessening the reduction in false memory as a result of negative valence. Furthermore, the fact that false memory differed for positive and negative moods even though arousal was equivalent in those groups indicates that valence contributed to differences in false memory as well, even when arousal was high. The likely explanation for this difference is reduced recollection rejection for subjects in positive moods.

The reason for the increased false memory in the negative mood group was related to increased phantom recollection and familiarity. In the first two studies, phantom recollection did not differ as a function of mood, indicating that increased phantom recollection is related to increased arousal. The increase in recollection rejection and identity judgment for negative moods observed in the first two studies did not occur here, again suggesting that elevating arousal changes the nature of processing in negative moods. In particular, the difference in parameter results across the three studies suggests that, consistent with predictions of FTT, arousal reduces verbatim processing. This proposition will be tested more directly in Experiment 4.

It should also be emphasized that the mood effect and its corresponding processes were isolated to short lists. Familiarity was consistently high across all three moods for long lists, and phantom recollection was highest for positive moods. The explanation for familiarity is clear: longer lists produce ceiling levels of familiarity, but shorter lists, which do not invoke familiarity to begin with, will have their familiarity levels raised by higher levels of arousal. As for phantom recollection, it may be particularly high for subjects in positive moods, for long lists, because of the effect of gist processing in positive moods combining with the strong gist produced by long lists, resulting in strong gist traces.

The effect of repeated testing on CDs was generally similar to the effect in Experiment 2, where repeated testing increased familiarity. Here, repeated testing increased both familiarity and phantom recollection (gist processes) for subjects in negative and neutral moods. Although this strong gist effect is a somewhat puzzling result for neutral moods, increased phantom recollection for negative moods is consistent with an arousal-based explanation. If higher arousal impairs storage of verbatim traces, then testing items immediately may lead to later confusing them with studied items and may result in higher levels of CD acceptance when they are tested for a second time. Whereas the repeated testing effect is driven by familiarity when arousal is low, heightened arousal may lead to even stronger feelings that items have been actually presented, resulting in phantom recollection.

The results of Experiment 3 provide evidence that arousal contributes to false memory, but that valence effects may remain even when arousal is elevated. The story so far is therefore that valence can affect false memory on its own, when arousal is controlled at a low level, and additionally that arousal can contribute to valence effects on false memory. However, the mood groups used in the present study do not allow for clear conclusions about the effects of arousal on false memory without any potential role of valence. The final piece of the puzzle will be to determine whether arousal affects false memory on its own, and in what way, when valence levels are controlled. This question was addressed in the final study.

CHAPTER 6

EXPERIMENT 4: EFFECT OF AROUSAL ON TRUE AND FALSE MEMORY FOR NEUTRAL WORDS

The mood manipulation used in Experiment 3 created moderate to high levels of arousal that were equivalent across the positive and negative mood groups. Although there was evidence that higher levels of arousal increased false memory, valence impacted false memory as well. To directly assess differences in arousal, while controlling for valence, it is necessary to compare three groups that differ in arousal but not in valence. Mirandola and Toffalini (2016) demonstrated that predetermined mood groups are not necessarily the best way to measure mood effects. Rather, they found that subjective ratings of arousal were better predictors of memory because of individual variation in arousal experienced within each group. There was not huge variability in valence experienced in response to the MIP, but arousal may be more subject to individual differences in responses to the MIP. Using the procedure Mirandola and Toffalini implemented, I similarly used subjective arousal ratings to compare three groups of subjects, with low, medium, and high arousal at the time of encoding. In other words, Experiment 4 involved a reclassification of the subjects from Experiment 3 into new arousal groups (rather than new data from a new group of subjects).

Method

Subjects from Experiment 3 were assigned to one of three arousal groups based on their post-induction subjective arousal ratings.

Results

Mood Manipulation Check

ANOVAs were conducted to confirm that the three new mood groups resulted in a) reliable arousal differences between all three groups and b) equivalent valence across all three groups. Post-manipulation arousal ratings were highest for the high group (M = 6.42, SD = .70), followed by the medium group (M = 4.51, SD = .51), which was followed by the low group (M = 2.85, SD = .56), $F(2,97) = 21.79$, $p < .05$. Post-manipulation valence ratings were statistically equivalent (M_{low} = 4.68, M_{medium} = 4.40, M_{high} = 5.18), $F(2,97) = 1.49$, *ns*.

Bias Correction

The effects of conjoint recognition question and mood on unrelated distractor acceptance were identified in order to determine the proper method of bias correction. Two parallel 3 (question: V, G, VG) x 3 (arousal: low, medium, high) analyses of variance (ANOVAs) were performed on the acceptance rates for unrelated distractors, separately for the immediate and the delayed tests. Arousal did not impact URD acceptance on either test.

Analysis of Standard True and False Memory

False memory was higher for subjects with high arousal (M = .46) compared to those with low arousal (M = .36) on the immediate test (false memory for medium arousal did not differ reliably from the other two groups), $F(2,95) = 3.26$, $p < .05$. This result suggests that arousal is sufficient to produce effects on false memory, and valence differences are not necessary to detect mood effects on false memory. There were no other reliable effects.

Main ANOVA Results

Four separate 3 (question: V, G, VG) x 3 (arousal: low, medium, high) x 2 (list length: short, long) ANOVAs were conducted on target and CD bias-corrected acceptance rates, separately for the immediate and delayed memory tests, with question and length as within-subjects factors and arousal as the between-subjects factor. Additional ANOVAs were conducted to measure forgetting and prior testing effects, as in the first three experiments. The results described below only include effects of arousal, as the other effects were consistent with the preceding studies.

Immediate test. There were no arousal effects for targets or CDs on the immediate test when all three instructional conditions were included in the analysis. This was because arousal influenced CD acceptance in the V condition as described in the previous section, but differences in CD acceptance were no longer reliable when the G and VG conditions were considered.

Delayed test. There was an arousal x length interaction for targets, $F(2,95) = 3.98$, $p < .05$. As shown in Figure 6.1, this interaction occurred because medium and high arousal improved target acceptance compared to low arousal, but only for short lists.

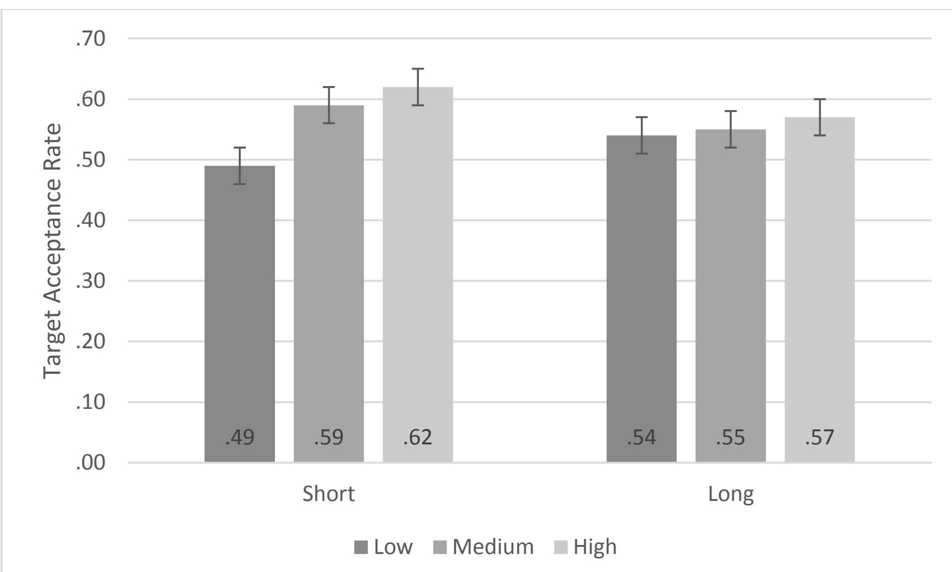


Figure 6.1. Target acceptance rate by length and arousal on the delayed test (Experiment 4).

There was a three-way interaction between arousal, length, and prior testing for CDs, $F(2,84) = 4.36$, $p < .05$. Prior testing had the greatest effect on subjects with low arousal, increasing CD acceptance for both list lengths (see Figure 6.2). Repeated testing did not impact subjects with high arousal, and impacted subjects with medium arousal by only increasing acceptance of CDs from short lists.

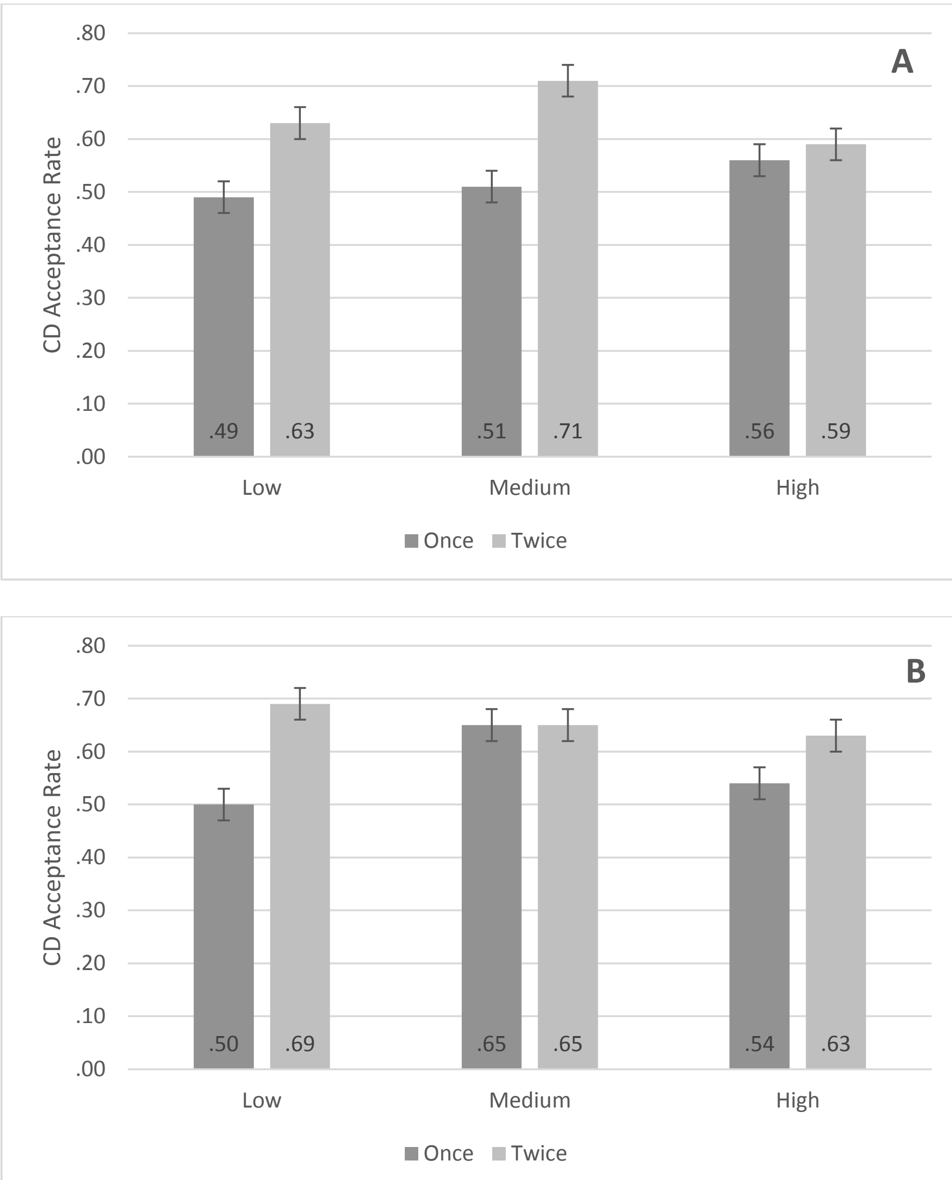


Figure 6.2. CD acceptance rate by arousal and prior testing for short lists (Panel A) and long lists (Panel B) on the delayed test (Experiment 4).

Finally, the effect of forgetting depicted in Figure 6.3 depended on arousal and CR instruction, $F(4,188) = 3.01$, $p < .05$. Subjects at all levels of arousal experienced forgetting in the V condition but not the G condition; subjects with low and medium arousal experienced forgetting in the VG condition but subjects with high arousal did not. Furthermore, CD acceptance was higher in the V than the G condition on the immediate test for subjects with medium and high arousal. On the delayed test, however, CD acceptance in the V condition had declined and became equivalent to G condition acceptance at all arousal levels.

Summary. High arousal increased false memory compared to low arousal on the immediate test. When analyses were extended to all instructional conditions, arousal did not impact CD acceptance on the immediate test, nor did it affect target acceptance. On the delayed test, however, medium arousal levels increased acceptance of targets from short lists compared to low arousal. Arousal did not impact CD acceptance overall on the delayed test, but it was affected by prior testing and forgetting. When arousal was high, prior testing had no effect, but when arousal was low prior testing increased CD acceptance. At moderate arousal levels, prior testing only increased CD acceptance for short lists. Subjects at all levels of arousal experienced forgetting in the V condition but did not experience forgetting in the G condition; only subjects with medium and low arousal experienced forgetting in the VG condition.

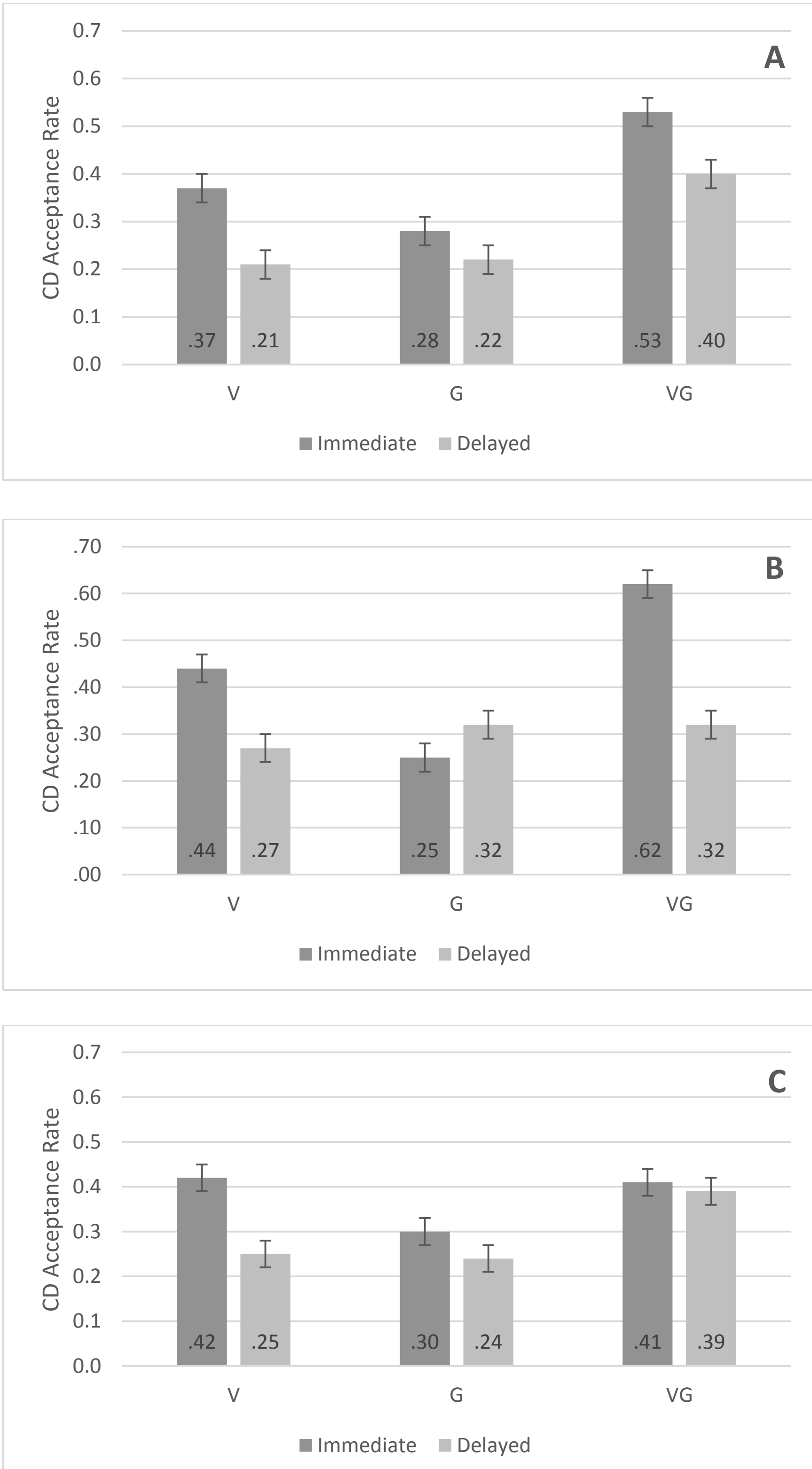


Figure 6.3. Forgetting of CDs by CD instruction for low (panel A), medium (panel B), and high (panel C) arousal moods in Experiment 4.

Model Fits

The model's fit was evaluated by computing likelihood ratio tests in the same manner it was done for the previous three studies. The G^2 values for all of the conditions were below 3.84, indicating that the model provided a statistically acceptable estimation of the data.

Parameter Estimates

High arousal increased standard false memory on the immediate test by increasing phantom recollection compared to low arousal, as shown in Table 6.1. This parameter difference was no longer reliable on the delayed test, consistent with the fact that arousal did not impact false memory on the delayed test. For targets on the delayed test, medium arousal levels increased acceptance compared to low arousal for short lists because of increased familiarity (identity judgment estimates were equivalent).

For CDs on the delayed test, list length affected whether prior testing increased CD acceptance for subjects with medium arousal levels: prior testing did not increase CD acceptance for long lists, and familiarity levels were equivalent regardless of prior testing (see Table 6.2). On the other hand, prior testing did increase CD acceptance for short lists, and familiarity was higher when items from short lists had been previously tested. Familiarity was high to begin with for long lists but not short lists. At low arousal, prior testing increased CD acceptance because of increased familiarity and phantom recollection and reduced recollection rejection. This pattern suggests that subjects in medium-aroused moods are more sensitive to differences in gist retrieval as a function of the number of items presented compared with low- and high-arousal subjects, where the effect of prior testing was consistent across list length.

Finally, forgetting reduced CD acceptance in the V condition for subjects with medium and high arousal levels because those subjects had elevated levels of phantom recollection on the immediate test but not after the delay, and they also experienced a decline in familiarity over time. This result suggests that higher arousal leads to increased phantom recollection but that the effect lessens over time. Forgetting in the V condition occurred for subjects with low arousal as well, but was driven by reduced familiarity and increased recollection rejection over time.

Table 6.1

Experiment 4 Parameter Estimates

	R_T	F_T	P_D	R_D	F_D	β_V	β_G	β_{VG}
Immediate Test								
Low	.25	.33	.22	.06	.42	.22	.44	.33
Medium	.21	.25	.29	.10	.40	.33	.33	.27
High	.29	.26	.38	.14	.40	.30	.40	.30
Short Low	.25	.26	.23	.24	.31	.24	.44	.30
Short Medium	.20	.28	.27	.09	.39	.32	.34	.26
Short High	.33	.19	.36	.11	.41	.30	.40	.30
Long Low	.37	.30	.26	.03	.46	.27	.44	.27
Long Medium	.25	.19	.34	.12	.42	.33	.33	.27
Long High	.26	.31	.43	.18	.40	.30	.40	.30
Delayed Test								
Short Low	.17	.09	.25	.16	.22	.38	.36	.38
Short Medium	.17	.17	.40	.30	.00	.37	.36	.29
Short High	.22	.32	.31	.21	.20	.42	.40	.42
Long Low	.13	.25	.37	.15	.36	.38	.36	.38
Long Medium	.17	.20	.18	.07	.24	.36	.36	.29
Long High	.18	.25	.31	.19	.33	.42	.40	.42
Low	.14	.18	.31	.16	.28	.38	.36	.38
Medium	.18	.18	.28	.18	.15	.36	.36	.29
High	.20	.28	.31	.20	.27	.42	.40	.42
Combined Across Both Tests								
Low	.20	.21	.29	.14	.33	.35	.39	.35
Medium	.15	.21	.28	.15	.24	.35	.35	.29
High	.24	.27	.34	.19	.31	.39	.39	.38

Table 6.2

Experiment 4 Prior Testing Parameter Estimates

	R_T	F_T	P_D	R_D	F_D	β_V	β_G	β_{VG}
				Not Previously Tested				
Low	.10	.03	.17	.19	.21	.38	.36	.38
Medium	.14	.06	.23	.16	.01	.36	.36	.29
High	.13	.09	.27	.23	.12	.42	.40	.42
Short Low	.14	.00	.06	.08	.16	.36	.36	.38
Short Medium	.11	.05	.22	.20	.00	.35	.35	.30
Short High	.17	.10	.18	.18	.07	.41	.41	.42
Long Low	.02	.15	.16	.21	.37	.38	.38	.38
Long Medium	.16	.08	.11	.00	.24	.36	.36	.29
Long High	.11	.07	.28	.21	.25	.41	.41	.42
				Previously Tested				
Low	.18	.26	.42	.12	.27	.38	.36	.38
Medium	.21	.28	.32	.20	.29	.36	.36	.29
High	.16	.32	.34	.17	.42	.42	.40	.42
Short Low	.14	.22	.35	.10	.27	.38	.38	.38
Short Medium	.22	.27	.21	.17	.43	.36	.36	.29
Short High	.14	.36	.30	.10	.46	.41	.41	.42
Long Low	.24	.30	.55	.17	.29	.38	.38	.38
Long Medium	.20	.29	.35	.20	.19	.36	.36	.29
Long High	.16	.29	.39	.22	.40	.41	.41	.42

Discussion

Experiment 4 addressed the question of arousal's effect on true and false memory when valence was held constant. When subjects were grouped based on low, medium, and high arousal levels, there were no valence differences among the three groups. The major new finding from this analysis method was that phantom recollection increased as arousal increased, which led to increased levels of false memory for subjects with higher levels of arousal. This finding is consistent with prior research showing that increasing subjects' arousal increases false memory (Corson & Verrier, 2007). This study is the first, however, to identify a mechanism for this effect—i.e., phantom recollection.

Phantom recollection has been connected to emotional content in prior work. Brainerd, Stein, et al. (2008) showed that although false memories were less likely to occur for positive words compared to negative and neutral words, when false memories did occur for positive words they were driven by phantom recollection, whereas negative and neutral false memories were driven by familiarity. On the other hand, emotional pictures reduced phantom recollection compared to neutral pictures in Bookbinder and Brainerd's (2017) design. Emotional content effects on phantom recollection may be sensitive to the type of study materials used or other methodological factors.

False memories accompanied by high levels of confidence may be similar to false memories driven by phantom recollection because of the strong feeling that a CD has been previously experienced. Indeed, false memories in high-arousing moods were accompanied by higher confidence ratings in Corson and Verrier's (2007) design, which is consistent with the present results. Remember judgments involve a similar feeling of item presentation accompanied by memory for details. Mirandola and Toffalini (2016) found these judgments to be reduced for positive and negative compared to neutral moods, but did not provide an analysis of remember/know judgments with respect to arousal. Clearly, more research is needed to clarify the connection between emotional moods and phantom recollection-driven false memory.

Repeated testing may increase phantom recollection; Roediger and McDermott (1995) showed that false memories accompanied by remember (as opposed to know) judgments were higher when items had been previously tested. The present data suggest that the mechanism for the repeated testing effect may depend on subjects' arousal levels. When arousal is low, subjects are more likely to accept previously tested items because of increased phantom recollection and familiarity and reduced recollection rejection. When arousal is only slightly higher, however, the repeated testing effect is driven by familiarity alone. Interestingly, the repeated testing effect disappeared when arousal was high. Considered in connection with the overall effect of arousal on false memory, one explanation for this pattern is that lower levels of false memory at lower arousal may be strengthened by items' presentation during a memory test, whereas the already high levels of false memory due to higher arousal will not be strengthened, because those subjects are already experiencing phantom recollection of presented items.

Similarly, the present results suggest that the processes driving forgetting of false memories (i.e., CD acceptance in the V condition) over time are different depending on subjects' arousal. When arousal is low, this forgetting is driven by a decline in familiarity and an increase in recollection rejection, but when arousal was high forgetting was driven by reduced phantom recollection. Rather than increasing over time, per se, it is possible that recollection rejection was inhibited on the immediate test for these subjects. Recollection rejection was particularly low on the immediate test, which implies that either memory for the surface features of targets was poor or CDs were not adequate retrieval cues for verbatim traces of targets (see Brainerd, Reyna, Wright, & Mojardin, 2003). If verbatim traces are affected by arousal according to the Yerkes-Dodson law (Yerkes & Dodson, 1908), access to these traces might be limited when arousal is particularly low, and may peak at a moderate arousal level. Furthermore, inspection of the parameter estimates in Table 6.1 reveals that the low recollection rejection for subjects in low arousing moods was driven by long lists: recollection rejection was at a more typical level (.24) for short lists but was only .03 for long lists. Perhaps the presentation of many targets strengthened gist traces to the point that surface details became fuzzy, particularly for subjects with low arousal. Again, however, it should be reiterated that the present study is the first to measure retrieval processes associated with arousal and further replications are needed before drawing definitive conclusions.

CHAPTER 7

GENERAL DISCUSSION

In this dissertation I explored the relationships between mood valence and arousal—separately—and true and false memory. The goal of these studies was to determine whether valence and arousal can have independent effects on false memory and to identify the retrieval processes involved in these effects.

Summary of Methodology

Through three pilot tests I developed two new MIPs that were able to successfully alter subjects' moods when administered online. The first MIP was able to alter subjects' valence while maintaining a low and equivalent arousal level across the three valence groups. The second MIP elevated the arousal level of the subjects in the positive and negative groups so that their arousal levels were equivalent to each other but higher than the neutral group, with valence still differing across the three groups. In Experiment 4, subjects were grouped instead by arousal, with valence being equal across low-, medium-, and high-arousal groups.

Experiment 1 tested the prediction that mood valence could impact false memory for 10-item DRM lists when mood was induced prior to the study phase. A mood manipulation check ensured that subjects' moods were successfully manipulated and differed in the intended way. Subjects in positive, negative, and neutral moods with low arousal completed immediate and two-day delayed conjoint recognition tests. The immediate recognition test was comprised of half of the test items and the delayed test was comprised of the items tested on the immediate test as well as the second half. The data were analyzed first using ANOVAs to identify effects on standard true and false memory. Then, the data from all three instructional conditions were analyzed and effects of the mood manipulation as well as the other factors were identified.

Finally, model fits, parameter estimates, and parameter differences tests were computed to identify the processes underlying the ANOVA results.

Experiment 2 followed the same general procedure as Experiment 1 and used the same mood manipulation. The primary difference was that instead of 10-item DRM lists, half of the lists contained 6 items and half of the lists contained 14 items. Experiment 3 followed the same procedure as Experiment 2, including the two list lengths, but the primary difference was in the MIP used. Instead of the low-arousal MIP, Experiment 3 used the high-arousal MIP in which the positive and negative mood groups had higher arousal levels than the neutral group.

Finally, in Experiment 4 the data from Experiment 3 were re-analyzed by grouping subjects based on their self-reported levels of arousal following the recommendation of Mirandola and Toffalini (2016). This restructuring resulted in three groups of subjects, with low, medium, and high arousal levels but equivalent valence levels on average. The data were analyzed in the same manner as the first three experiments with a focus on new effects of arousal and their accompanying retrieval processes.

Summary of Results

The Effect of Mood on False Memory

Negative moods reduced false memory compared to positive moods because of enhanced verbatim processing, confirming the prediction of Experiment 1. The effect was replicated in Experiment 2 with two list lengths, but the mechanism was different: Negative moods reduced false memory for long lists by increasing recollection rejection, but they reduced false memory for short lists by reducing familiarity. When arousal in the positive and negative mood groups was increased, in Experiment 3, negative moods still reduced false memory compared to positive moods, but now false memory was lowest for subjects in neutral moods and the differences among all three mood groups were reliable. This new pattern of results suggested that arousal increased false memory, but that valence effects remained. Increased false memory in high-arousal positive moods compared to negative moods was driven by reduced recollection rejection (i.e., a verbatim process), and increased false memory in negative moods (compared to neutral moods was driven by increased phantom recollection and familiarity (i.e., gist processes). These findings provided initial evidence that negative moods reduce false memory via verbatim processing and arousal increases false memory via both verbatim and gist processing. Experiment 4 provided a direct test of the prediction that arousal can affect false memory independent from valence, confirming that prediction and locating the mechanism primarily within gist processing, via increased phantom recollection. Together, these results provide a coherent story of mood-false memory effects: Negative moods reduce false memory compared to positive moods by emphasizing verbatim processing and high arousal moods increase false memory compared to low arousal moods by emphasizing gist processing.

Additional Manipulations

The retrieval processes that contribute to several other manipulations were identified, which served several purposes: a) to confirm the predictions that have been made about the way these manipulations dissociate verbatim and gist memory and b) to provide further support for the verbatim valence/gist arousal hypothesis through the way these manipulations interacted with mood valence and arousal.

List length. Experiment 2 revealed that false memory was higher for longer lists because of reduced recollection rejection and increased familiarity. When arousal was increased in Experiment 3 mood effects on false lists were isolated to short lists because familiarity was high for long lists across all moods. This effect is unsurprising because longer lists repeat the gist of the list more than short lists. Emotional moods with high arousal increased familiarity for short lists, where familiarity was low to begin with, because higher arousal increased gist retrieval.

Repeated testing. In Experiment 1 repeated testing increased familiarity and phantom recollection of CDs. Experiment 2 revealed that the repeated testing effect depended on list length because it increased CD acceptance for short lists more than long lists, again because longer lists provide stronger gist due to repeating the gist to subjects. In Experiments 3 and 4, repeated testing increased phantom recollection in negative and neutral moods but not positive moods, and in low arousal moods but not high arousal moods, because phantom recollection was already high when arousal was high.

Forgetting. Forgetting of CDs in Experiment 1 was a result of reduced phantom recollection over time, whereas forgetting of targets was due to reduced identity judgment. When the list length manipulation was introduced in Experiment 2 forgetting occurred more for short lists than long lists and was driven by reductions in familiarity for CDs. When arousal was high in Experiment 4, forgetting of CDs was driven by declines in phantom recollection whereas when arousal was low it was driven by declines in familiarity and increased recollection rejection. Furthermore, forgetting of both targets and CDs occurred more for short lists than long lists.

Theoretical Implications

The main goals of this dissertation were to determine the nature of valence and arousal effects on false memory with controlled methods, to identify the nature of those effects, and to identify the conditions under which false memory is most likely to occur. To achieve those goals I posed several main research questions.

Experiment 1 addressed the question of whether mood valence impacts false memory even when there is no potential influence of arousal. The finding that negative mood reduced false memory and positive moods increased false memory compared to neutral moods is consistent with predictions of both AAI and FTT. False memories should be more likely when subjects use relational processing or favor gist retrieval, whereas they should be less likely when subjects use item-specific processing or favor verbatim retrieval. Because arousal was controlled at a low level, this finding provides credence to the claims made previously that mood valence can—on its own—influence false memory (Storbeck & Clore, 2005, 2011). Estimation of the parameters of the CR model provided a more direct test of the processing explanation. Consistent with the verbatim hypothesis, negative moods increased identity judgment. However, as identity judgment is a verbatim process but does not have a direct impact on false memory, Experiment 2 was designed to further dissociate verbatim and gist retrieval and provide a more cohesive process explanation.

When two list lengths were used in Experiment 2, negative mood still reduced false memory but did so via different processes for short versus long lists, confirming that the list length manipulation dissociated verbatim from gist retrieval. Indeed, in Experiment 3, mood effects were isolated to short lists, where verbatim traces were stronger, providing further support for the hypothesis that negative moods reduce false memory via stronger verbatim processing.

Experiment 3 was designed to determine how mood valence effects on false memory would change when arousal was increased for positive and negative moods. The fact that the pattern of mood valence effects changed indicated that arousal had an impact on false memory as well. In Experiments 1 and 2, negative moods reduced false memory compared to neutral moods (although the only reliable difference was between negative and positive moods). In particular, the fact that now both positive and negative moods reliably increased false memory compared to neutral moods provided an important theoretical distinction. AAI predicts that increased arousal strengthens the current active processing style. If more false memories occur in positive moods than negative moods when arousal is low, this gap should expand when arousal is high, increasing the difference between valenced moods from neutral moods. On the other hand, FTT predicts that increased arousal would weaken verbatim memory traces, increasing false memory for subjects in negative moods. The results provide support for FTT's prediction but not AAI's prediction, because when arousal was increased false memory was higher in negative moods compared to neutral moods.

The results from Experiment 3, considered with the results from the first two experiments, provide context for interpreting prior research as well. The lack of control of arousal in early mood-false memory research made it difficult to conclude with certainty that mood valence was truly the only cause of false memory differences (Storbeck & Clore, 2005, 2011). Here, I found that a) when arousal was controlled the pattern of valence effects was the same as what was found previously and b) when arousal was increased the pattern of valence effects was different from what was found previously. Therefore, if prior research was suffering from arousal differences among the valence groups, one might expect results more similar to Experiment 3. The fact that their results were more similar to Experiment 1 supports their claim that it was indeed valence, and not arousal, that caused the observed differences in false memory. Of course, it is still not possible to know without any arousal data from those studies, and of course it is still best practice to include proper controls in the first place. That being said, I would argue that the current results aid in correct interpretation of early results.

Finally, Experiment 4 demonstrated that arousal influences false memory independent of valence. In particular, higher arousal levels led to higher levels of false memory because of increased phantom recollection. When subjects experienced higher arousal, they stored gist traces that were so strong that they thought CDs had been previously presented, and perhaps even had specific memories of their prior presentation. The arousal result is consistent with Corson and Verrier's (2007) claim that high arousal increases false memory and expands upon that result by identifying phantom recollection as the mechanism for that effect. That arousal could influence false memory independent from valence is consistent with predictions of FTT but not AAI. In AAI, arousal does not directly affect processing; rather, it influences whatever processing is currently being employed, encouraging that processing when arousal is high and attenuating it when arousal is low. AAI does not explain how arousal could influence false memory if neither item-specific nor relational processing is already being favored. On the other hand, in FTT arousal may influence verbatim or gist processing directly. The fact that high arousal enhanced phantom recollection is therefore consistent with FTT but not AAI, because high arousal is expected to reduce access to verbatim traces according to FTT.

Using the same method of grouping subjects based on self-reported arousal, Mirandola and Toffalini (2016) found the opposite—that arousal reduced false memory. Their finding is not inconsistent with the present research because of an important methodological difference. Mirandola and Toffalini manipulated mood after the study phase and before their recognition test, whereas I manipulated it prior to the study phase. Both AAI and FTT predict that mood influences false memory at encoding, by either affecting how targets are processes or memory traces of them are stored. Neither theory makes a direct prediction regarding how mood would affect false memory when manipulated after targets have already been presented.

Some might argue that the fact that there were no mood effects on false memory after a delay is evidence that induced moods need to be currently active in order to affect retrieval. However, the lack of delayed test results can easily be explained by the fact if mood valence affects verbatim traces more than gist traces, and verbatim traces decline more rapidly over time, mood effects may disappear. Indeed, forgetting occurred more in the V condition than in the G condition consistently in all four experiments because verbatim processes declined over time more so than gist processes, especially for targets. Forgetting also declined more for short lists than long lists, because more verbatim traces were stored for short lists and more gist traces were stored for long lists. However, forgetting occurred for subjects in all moods. Despite retention interval being a verbatim manipulation, it did not impact subjects in negative moods more than others. If it is true that forgetting of both accurate and false memories is not impacted by mood, that would be a positive finding in the context of real-life remembering, because negative moods would not promote more risk of inaccuracy due to forgetting.

On the other hand, repeated testing affected subjects in positive and negative moods differently. Although repeated testing increased both true and false memory for subjects in all moods, subjects in negative-arousing moods were more likely to have false memories driven by phantom recollection and lower levels of recollection rejection for items that had been tested previously. This difference is of high importance because it means that people in negative moods may be more confident in false memories that are based on earlier memory tests, compared to people in positive moods. Police interviews in which witnesses are asked similar questions repeatedly could lead people—especially those who are feeling fearful or otherwise negative during the reporting process—to state with certainty that they remember something that did not occur but that they were asked about earlier.

More generally, when arousal was low repeated testing increased false memory across all mood valences because it increased familiarity and sometimes phantom recollection. Although negative-arousing moods may put people at particular risk for illusory memories due to repeated testing, repeated testing will increase false memory via weaker gist traces for people in a variety of moods. This effect was particularly strong for short lists. Although longer lists increased false memory in general compared to short lists, short lists were more susceptible to the repeated testing effect, because they provided lower levels of familiarity to begin with. Therefore, if there is less gist-consistent information to be remembered, the act of repeatedly testing it may create more distortion by providing better gist, compared to repeatedly testing information that has a stronger gist to begin with.

There is other evidence for repeated testing increasing phantom recollection from other conjoint recognition research. Bookbinder and Brainerd (2017) found that an earlier memory test increased phantom recollection of emotional pictures. They also found that phantom recollection was higher on the immediate memory test and declined over time, which is somewhat consistent with the present research. I found that forgetting of CDs was driven primarily by familiarity, but sometimes by phantom recollection as well. In both designs, forgetting occurred more for targets than CDs because it occurred more for verbatim than gist memory traces.

The results of the four studies in this dissertation provide the following picture of the conditions most apt to produce false memories. Positive moods foster false memories more than negative moods and arousing moods create more false memories than non-arousing moods. Although negative moods can be protective against false memory, negative-arousing moods can promote distortion. This pattern of results supports FTT better than other mood theories. False memories may fade with time, but they are more robust over time compared to true memories. Repeatedly testing false memory items can lead to even stronger false memories, particularly for people in negative-aroused moods. To-be-remembered items that create a strong gist (through longer study lists) will lead to more false memories, but false memories for items with a weaker gist (short lists) can be strengthened through repeated testing. Forgetting of true memories is more likely for short lists, as well, which means that short lists may be more prone to distortion than longer lists over time.

Applications

Knowledge about the context that is most conducive to false memories may be useful outside the laboratory. This dissertation has provided some preliminary conditions under which false memories may be particularly strong. Importantly, some of those conditions may map onto the conditions in certain real-life situations when false memories can be dangerous. Negative, low-arousal moods may protect against memory distortion, potentially because it would be adaptive to remember the specific details of risky or threatening situations (Schwarz & Clore, 1983). However, the protective nature of negative moods apparently disappears when arousal becomes higher, at least in comparison to neutral moods. The fact that increased arousal may promote memory distortion especially in negative moods—and that that distortion is accompanied by phantom recollections of experiencing non-experienced events—is of particular concern when recalling traumatic events in therapy or reporting on witnessed crimes.

Although the applicability of the DRM paradigm to such real-life situations has been questioned (e.g., DePrince, Allard, Oh, & Freyd, 2004), Gallo (2010; 2013) reviewed several lines of evidence pointing to its relevance. For example, there is a correlation between autobiographical false memory and false memory in the DRM illusion (e.g., Platt, Lacey, Iobst,

& Finkelman, 1998). Additionally, brain damage that affects autobiographical memory can also affect false memories in the DRM paradigm (e.g., Schacter, 2012). The data show that in fact the DRM illusion is a good predictor of real-world memory performance despite the criticisms that have been put forth.

Others have questioned the generalizability of experimentally-induced moods, as well (Martin, 1990), whereas proponents of MIPs have attested to the reality of the emotions experienced in response to these procedures, especially when videos are used (Rottenberg, Ray, & Gross, 1995). One alternative would be to study differences in false memory in clinical populations with a variety of moods, which has clear importance, but introduces additional issues and confounding variables (Otgaar, Muris, Howe, & Merckelbach, 2017). Experimentally induced moods have the advantage of not confounding trait and state moods and allow for better control of valence and arousal (when the methodology permits, of course).

Limitations and Future research

Despite the advantages of an experimental MIP using videos, a limitation of this design may be the artificiality of the induced moods and their short time span. One avenue for future research could be the replication of this design with a MIP that provides a closer approximation of moods experienced in real life, perhaps with a particular focus on forensic applications. Such a mood induction would be limited by the types of emotions that are ethically acceptable to induce, but there might be a middle ground. Even without a forensic focus, it would be beneficial to replicate these effects using a more realistic mood induction; for example, Yang, Yang, Ceci, and Isen (2015) gave subjects candy as a positive mood induction. Despite the limitations of studying moods that cannot be experimentally controlled, it would also be of interest to identify retrieval processes involved in false memory in naturally-occurring clinical and non-clinical moods.

Concerning the generalizability of the DRM paradigm, studies using the same controlled mood induction within other false memory paradigms will be useful. There is evidence that false memory for words and pictures are affected differently by the emotional content of the stimuli (Bookbinder & Brainerd, 2016), so it is possible that mood may affect false memory for pictures differently. As there is little research on how mood affects picture memory, studies with controlled valence and arousal, and studies that measure retrieval processes, may provide new information. Similarly, there is yet little research on mood effects on implanted false memory in the misinformation paradigm. Use of this paradigm is another way to enhance the applicability of false memory results, but we have yet to fully identify mood's role in implanted false memory and the associated retrieval processes.

With a similar focus on enhancing the relevance of mood-false memory research it will be necessary to apply developmental considerations. MIPs have not often been applied to children, particularly in the context of memory, but researchers in other fields have succeeded at altering children's moods with consequential effects on information processing (Harper, Lemerise, & Caverly, 2010). We have begun to identify factors that make children more or less prone to false memory compared with adults, and emotional content may be one of those factors (Brainerd, Reyna, & Ceci, 2008). Mood may similarly tie into the developmental trajectory of false memory, which is important information to have both theoretically and because of children's essential role in legal proceedings and the debates surrounding it. A similar investigation should be done with the aging population as well, based on evidence that proneness to false memory may change with healthy aging (Koutstaal & Schacter, 1997) as well as the tendency to adopt certain moods or remember emotional information (Mather & Carstensen, 2005).

The negative-arousing moods induced in this research may be relevant to the law and other real-life remembering situations, but rarely is it the case that the information being reported is entirely neutral. Evidence of mood-congruent true and false memory has accumulated in recent years, but there are still unanswered questions. Some of those questions remain unanswered because these studies have not been immune to the issue of conflated valence and arousal seen in the rest of the emotion literature. Careful control of valence and arousal and identification of retrieval processes may begin to answer those questions.

In short, there remain many directions for future mood-false memory research. Many of them will provide information that can be applied outside of the laboratory and others will continue to aid in the refinement of theoretical conceptions of human emotion.

Concluding Comment

The debate about whether mood-dependent false memory is driven by valence or arousal may be answered with a compromise: Mood valence can affect false memory, but arousal can affect false memory just as well. At certain levels of valence and arousal, the result may be the perfect combination to create high levels of distortion that are accompanied by strong feelings that those distortions are, in fact, real. The knowledge that both valence and arousal contribute to false memory and that their impacts are due to different retrieval processes, as well as understanding of the other factors that make memory particularly unreliable, may be of high practical importance.

References

Adelman, J. S., & Estes, Z. (2013). Emotion and memory: A recognition advantage for positive and negative words independent of arousal. *Cognition*, *129*, 530–535.

Alba, J. W., & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin*, *93*, 203–231.

Anderson, A. K., Christoff, K., Panitz, D., De Rosa, E., & Gabrieli, J. D. E. (2003). Neural correlates of the automatic processing of threat facial signals. *Journal of Neuroscience*, *23*, 5627–5633.

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

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

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

Bower, G. H. (1981). Mood and memory. *The American Psychologist*, *36*, 129–148.

Bradley, M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*, 49–59.

Bradley, M. M., & Lang, P. J. (1999). *Affective Norms for English Words (ANEW): Stimuli, Instruction Manual and Affective Ratings*. *Technical Report C-1*. Gainsville, FL.

Bradley, M. M., & Lang, P. J. (2007). The International Affective Picture System (IAPS) in the study of emotion and attention. *Handbook of Emotion Elicitation and Assessment*, 29–46.

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

Experimental Child Psychology, *107*, 137–154.

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

Brainerd, C. J., & Reyna, V. F. (2007). Explaining developmental reversals in false memory. *Psychological Science*, *18*, 442–448.

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

Brainerd, C. J., Reyna, V. F., Wright, R., & Mojardin, A. H. (2003). Recollection rejection: False-memory editing in children and adults. *Psychological Review*, *110*, 762–784.

Brainerd, C. J., Stein, L. M., Silveira, R. A., Rohenkohl, G., & Reyna, V. F. (2008). How Does Negative Emotion Cause False Memories? *Psychological Science*, *19*, 919–925.

Brainerd, C. J., Wright, R., Reyna, V. F., & Mojardin, A. H. (2001). Conjoint recognition and phantom recollection. *Journal of Experimental Psychology: Learning Memory and Cognition*, *27*, 307–327.

Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition*, *5*, 73–99.

Budson, A. E., Todman, R. W., Chong, H., Adams, E. H., Kensinger, E. A., Krangel, T. S., & Wright, C. I. (2006). False recognition of emotional word lists in aging and Alzheimer disease. *Cognitive and Behavioral Neurology*, *19*, 71–78.

Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1997). Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. *Personality and Social Psychology Review*, *1*, 3–25.

Cahill, L., Gorski, L., & Le, K. (2003). Enhanced human memory consolidation with post-learning stress: interaction with the degree of arousal at encoding. *Learning & Memory, 10*,

270–274.

Choi, H.-Y., Kensinger, E. A., & Rajaram, S. (2013). Emotional content enhances true but not false memory for categorized stimuli. *Memory & Cognition*, *41*, 403–415.

Clore, G. L., & Storbeck, J. (2006). Affect as information about liking, efficacy, and importance. *Affect in Social Thinking and Behavior*. New York, NY, US: Psychology Press.

Colibazzi, T., Posner, J., Wang, Z., Gorman, D., Gerber, A., Yu, S., … Peterson, B. S. (2010). Neural systems subserving valence and arousal during the experience of induced emotions. *Emotion*, *10*, 377–389.

Corson, Y., & Verrier, N. (2007). Emotions and false memories: Valence or arousal? *Psychological Science*, *18*, 208–211.

Dan-Glauser, E. S., & Scherer, K. R. (2011). The Geneva affective picture database (GAPED): a new 730-picture database focusing on valence and normative significance. *Behavior Research Methods*, *43*, 468–77.

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

Dehon, H., Larøi, F., & Van der Linden, M. (2010). Affective valence influences participant's susceptibility to false memories and illusory recollection. *Emotion*, *10*, 627–39.

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

DePrince, A. P., Allard, C. B., Oh, H., & Freyd, J. (2004). What's in a name for memory errors? Implications and ethical issues arising from the use of the term "false memory" for errors in memory for details. *Ethics & Behavior*, *14*, 255–295.

Dewhurst, S. A., Pursglove, R. C., & Lewis, C. (2007). Story contexts increase susceptibility to the DRM illusion in 5-year-olds. *Developmental Science*, *10*, 374–378.

Dutton, D. G., & Aron, A. P. (1974). Some evidence for heightened sexual attraction under conditions of high anxiety. *Journal of Personality and Social Psychology*, *30*, 510–517.

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review*, *66*, 183–201.

Eerola, T., & Vuoskoski, J. K. (2010). A comparison of the discrete and dimensional models of emotion in music. *Psychology of Music*, *39*, 18–49.

Emery, L., Hess, T. M., & Elliot, T. (2012). The illusion of the positive: The impact of natural and induced mood on older adults' false recall. *Aging, Neuropsychology, and Cognition*, *19*, 677–698.

Gabert-Quillen, C. A., Bartolini, E. E., Abravanel, B. T., & Sanislow, C. A. (2015). Ratings for emotion film clips. *Behavior Research Methods*, *47*, 773–787.

Gallo, D. A. (2010). False memories and fantastic beliefs: 15 years of the DRM illusion. *Memory & Cognition*, *38*, 833–848.

Gallo, D. A. (2013). *Associative illusions of memory: false memory research in DRM and related tasks*. New York, NY: Psychology Press.

Gallo, D. A., McDermott, K. B., & Percer, J. M. (2001). Modality effects in false recall and false recognition. *Cognition*, *27*, 339–353.

Gomes, C. F., Brainerd, C. J., & Stein, L. M. (2013). Effects of emotional valence and arousal on recollective and nonrecollective recall. *Journal of Experimental Psychology: Learning Memory and Cognition*, *39*, 663–677.

Gross, J. J., & Levenson, R. W. (1995). Emotion elicitation using films. *Cognition & Emotion*, *9*,

87–108.

Hamann, S. (2001). Cognitive and neural mechanisms of emotional memory. *Trends in Cognitive Sciences*, *5*, 394–400.

Harper, B. D., Lemerise, E. A., & Caverly, S. L. (2010). The Effect of Induced Mood on Children's Social Information Processing: Goal Clarification and Response Decision. *Journal of Abnormal Child Psychology*, *38,* 575–586.

Heuer, F., & Reisberg, D. (1990). Vivid memories of emotional events: The accuracy of remembered minutiae. *Memory & Cognition*, *18*, 496–506.

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

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

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

Hurlemann, R., Hawallek, B., Matusch, A., Kolsch, H., Wollersen, H., Madea, B., … Dolan, R. J. (2005). Noradrenergic modulation of emotion-induced forgetting and remembering. *Journal of Neuroscience*, *25*, 6343–6349.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*, 513–541.

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

Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory & Cognition*, *31*,

1169–1180.

Kensinger, E. A., Garoff-Eaton, R. J., & Schacter, D. L. (2007). Effects of emotion on memory specificity: Memory trade-offs elicited by negative visually arousing stimuli. *Journal of Memory and Language*, *56*, 575–591.

Kensinger, E. A., O'Brien, J. L., Swanberg, K., Garoff-Eaton, R. J., & Schacter, D. L. (2007). The effects of emotional content on reality-monitoring performance in young and older adults. *Psychology and Aging*, *22*, 752–764.

Knott, L. M., & Thorley, C. (2013). Mood-congruent false memories persist over time. *Cognition & Emotion*, *28*, 903–912.

Knott, L. M., Threadgold, E., & Howe, M. L. (2014). Negative mood state impairs false memory priming when problem-solving. *Journal of Cognitive Psychology*, *26*, 580–587.

Koutstaal, W., & Schacter, D. L. (1997). Gist-based false recognition of pictures in older and younger adults. *Journal of Memory and Language*, *37*, 555–583.

Lampinen, J. M., & Schwartz, R. M. (2000). The impersistence of false memory persistence. *Memory*, *8*, 393–400.

Lang, P., Bradley, M., & Cuthbert, B. (2005). *International Affective Picture System (IAPS): Digitized Photographs, Instruction Manual and Affective Ratings. Technical Report A-6. Report A-6*. Gainsville, FL.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1998). Emotion, motivation, and anxiety: Brain mechanisms and psychophysiology. *Biological Psychiatry*, *44*, 1248–1263.

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, *30*, 261–273.

Larsen, J. T. (2017). Holes in the case for mixed emotions. *Emotion Review*, *9*, 118–123.

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

Martin, M. (1990). On the induction of mood. *Clinical Psychology Review*, *10*, 669–697.

Mather, M. (2007). Emotional arousal and memory binding: An object based framework. *Perspectives on Psychological Science, 2*, 33–52.

Mather, M., & Carstensen, L. L. (2005). Aging and motivated cognition: The positivity effect in attention and memory. *Trends in Cognitive Sciences*, *9,* 496–502.

Mattek, A. M., Wolford, G. L., & Whalen, P. J. (2017). A mathematical model captures the structure of subjective affect. *Perspectives on Psychological Science*, *12*, 508–526.

Mayer, J. D., Allen, J. P., & Beauregard, K. (1995). Mood inductions for four specific moods: A procedure employing guided imagery vignettes with music. *Journal of Mental Imagery, 19*, 133–150.

Mayer, J. D., & Gaschke, Y. N. (1988). The experience and meta-experience of mood. *Journal of Personality and Social Psychology*, *55,* 102–111.

McDermott, K. B. (1996). The persistence of false memories in list recall. *Journal of Memory and Language*, *35*, 212–230.

McDermott, K. B. (2006). Paradoxical effects of testing: Repeated retrieval attempts enhance the likelihood of later accurate and false recall. *Memory & Cognition*, *34,* 261–267.

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*, 160–176.

Mirandola, C., & Toffalini, E. (2016). Arousal-but not valence-reduces false memories at retrieval. *PLoS ONE*.

Otgaar, H., Muris, P., Howe, M. L., & Merckelbach, H. (2017). What drives false memories in psychopathology? A case for associative activation. *Clinical Psychological Science*, 1–22.

Phelps, E. A., Ling, S., & Carrasco, M. (2006). Emotion facilitates perception and potentiates the perceptual benefits of attention. *Psychological Science*, *17*, 292–299.

Platt, R. D., Lacey, S. C., Iobst, A. D., & Finkelman, D. (1998). Absorption, dissociation, and fantasy-proneness as predictors of memory distortion in autobiographical and laboratory-generated memories. *Applied Cognitive Psychology*, *12*, S77–S89.

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

Qualtrics. (2017). Qualtrics. Provo, UT.

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

Reyna, V. F., & Kiernan, B. (1994). Development of gist versus verbatim memory in sentence recognition: Effects of lexical familiarity, semantic content, encoding instructions, and retention interval. *Developmental Psychology*, *30,* 178–191.

Rhodes, M. G., & Anastasi, J. S. (2000). The effects of a levels-of-processing manipulation on false recall. *Psychonomic Bulletin & Review*, *7*, 158–162.

Robinson, K. J., & Roediger, H. L. (1997). Associative processes in false recall and false recognition. *Psychological Science*, *8,* 231–237.

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.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 95–115). Washington, DC, US: American Psychological Association.

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

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,* 385–407.

Rottenberg, J., Ray, R. D., & Gross, J. J. (1995). Emotion elicitation using films. *Cognition and Emotion*, *9*, 87–108.

Ruci, L., Tomes, J. L., & Zelenski, J. M. (2009). Mood-congruent false memories in the DRM paradigm. *Cognition and Emotion*, *23*, 1153–1165.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*, 1161–1178.

Russell, J. A. (1991). Culture and the categorization of emotions. *Psychological Bulletin*, *110,* 426–450.

Russell, J. A., & Carroll, J. M. (1999). On the bipolarity of positive and negative affect. *Psychological Bulletin*, *125*, 3–30.

Schacter, D. L. (2012). Constructive memory: Past and future. *Dialogues in Clinical Neuroscience*, *1114*, 7–18.

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

Schaefer, A., Nils, F., Sanchez, X., & Philippot, P. (2010). Assessing the effectiveness of a large database of emotion-eliciting films: A new tool for emotion researchers. *Cognition & Emotion*, *24*, 1153–1172.

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, *45*, 513–523.

Seamon, J. G., Luo, C. R., Shulman, E. P., Toner, S. K., & Caglar, S. (2002). False memories are hard to inhibit: Differential effects of directed forgetting on accurate and false recall in the DRM procedure. *Memory*, *10*, 225–237.

Semmler, C., & Brewer, N. (2002). Effects of mood and emotion on juror processing and judgments. *Behavioral Sciences & the Law*, *20*, 423–436.

Shaw, J., & Porter, S. (2015). Constructing rich false memories of committing crime. *Psychological Science*, *26*, 291–301.

Smith, R. E., & Hunt, R. R. (1998). Presentation modality affects false memory. *Psychonomic Bulletin & Review*, *5*, 710–715.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: applications to dementia and amnesia. *Journal of Experimental Psychology. General*, *117*, 34–50.

Spanos, N. P. (1996). *Multiple identities & false memories: A sociocognitive perspective.* Washington: American Psychological Association.

Stadler, M. A., Roediger, H. L., Ill, & Mcdermott, K. B. (1999). Norms for word lists that create false memories. *Memory & Cognition*, *27*, 494–500.

Storbeck, J. (2013). Negative affect promotes encoding of and memory for details at the expense of the gist: affect, encoding, and false memories. *Cognition & Emotion*, *27*, 800–819.

Storbeck, J., & Clore, G. L. (2005). With sadness comes accuracy; with happiness, false memory: Mood and the false memory effect. *Psychological Science*, *16*, 785–791.

Storbeck, J., & Clore, G. L. (2008). Affective arousal as information: How affective arousal influences judgments, learning, and memory. *Social and Personality Psychology Compass*, *2*, 1824–1843.

Storbeck, J., & Clore, G. L. (2011). Affect influences false memories at encoding: evidence from recognition data. *Emotion*, *11*, 981–989.

Storbeck, J., & Clore, G. L. (2011). Affect influences false memories at encoding: Evidence from recognition data. *Emotion*, *11*, 981–989.

Sugrue, K., & Hayne, H. (2006). False memories produced by children and adults in the DRM paradigm. *Applied Cognitive Psychology*, *20*, 625–631.

Talarico, J. M., & Rubin, D. C. (2003). Flashbulb memories. *Psychological Science*, *14*, 455–461.

Talmi, D., Luk, B. T. C., McGarry, L. M., & Moscovitch, M. (2007). The contribution of relatedness and distinctiveness to emotionally-enhanced memory. *Journal of Memory and Language*, *56*, 555–574.

Talmi, D., & Moscovitch, M. (2004). Can semantic relatedness explain the enhancement of memory for emotional words? *Memory & Cognition*, *32*, 742–751.

Tellegen, A., Watson, D., & Clark, L. (1999). On the dimensional and hierarchical structure of affect. *Psychological Science*, *10*, 297–303.

Thapar, A., & McDermott, K. B. (2001). False recall and false recognition induced by presentation of associated words: Effects of retention interval and level of processing. *Memory & Cognition*, *29*, 424–432.

Toglia, M. P., Neuschatz, J. S., & Goodwin, K. A. (1999). Recall accuracy and illusory memories: when more is less. *Memory*, *7*, 233–256.

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

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

Experimental Psychology, *70*, 122–129.

Van Damme, I. (2013). Mood and the DRM paradigm: An investigation of the effects of valence and arousal on false memory. *Quarterly Journal of Experimental Psychology*, *66*, 1060–1081.

Verheyen, C., & Göritz, A. S. (2009). Plain Texts as an Online Mood-Induction Procedure. *Social Psychology*, *40*, 6–15.

Watson, D., Clark, L., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS Scales. *Journal of Personality and Social Psychology*, *54*, 1063–1070.

Yang, H., Yang, S., Ceci, S. J., & Isen, A. M. (2015). Positive affect facilitates the effect of a warning on false memory in the DRM paradigm. *The Journal of Positive Psychology*, *10*, 196–206.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, *18*, 459–482.

Yonelinas, A. P. (1997). Recognition memory ROCs for item and associative information: The contribution of recollection and familiarity. *Memory & Cognition*, *25*, 747–763.

Yonelinas, A. P. (2002). The Nature of Recollection and Familiarity: A Review of 30 Years of Research. *Journal of Memory and Language*, *46*, 441–517.

Appendix A

Experiment 1 Acceptance Probabilities

Unrelated Distractors – Immediate Test

		M	*SD*
	Negative	.21	.27
V	Neutral	.24	.28
	Positive	.26	.27
	Total	.24	.27
	Negative	.31	.30
G	Neutral	.37	.34
	Positive	.47	.31
	Total	.38	.33
	Negative	.16	.27
VG	Neutral	.27	.31
	Positive	.33	.30
	Total	.25	.30
	Negative	.23	.36
Total	Neutral	.29	.29
	Positive	.36	.38

Unrelated Distractors – Delayed Test

		M	*SD*	*M*	*SD*	*M*	*SD*
		Tested		Not Tested		Total	
V	Negative	.51	.34	.37	.29	.42	.32
	Neutral	.42	.35	.39	.31	.40	.34
	Positive	.52	.27	.56	.35	.54	.31
	Total	.47	.33	.42	.32	.44	.32
G	Negative	.32	.35	.31	.34	.31	.35
	Neutral	.33	.32	.22	.27	.27	.30
	Positive	.41	.36	.35	.32	.38	.34
	Total	.34	.34	.28	.31	.31	.33
VG	Negative	.44	.33	.36	.39	.40	.36
	Neutral	.42	.38	.33	.31	.41	.35
	Positive	.37	.31	.35	.29	.44	.30
	Total	.42	.35	.34	.33	.41	.32
Total	Negative	.42	.34	.35	.34	.34	.34
	Neutral	.41	.35	.32	.29	.32	.32
	Positive	.41	.30	.42	.32	.41	.31

Targets – Immediate Test

		M	*SD*
V	Negative	.30	.23
	Neutral	.25	.25
	Positive	.25	.25
	Total	.27	.24
G	Negative	.17	.21
	Neutral	.24	.23
	Positive	.10	.16
	Total	.18	.21
VG	Negative	.54	.27

	Neutral	.43	.36
	Positive	.38	.30
	Total	.45	.33
	Negative	.33	.23
Total	Neutral	.30	.29
	Positive	.24	.24

Targets – Delayed Test

Valence	Question	Times Tested	*M*	*SD*
Negative	V	Once	.13	.15
		Twice	.25	.17
		Total	.19	.16
	G	Once	.19	.22
		Twice	.19	.25
		Total	.19	.23
	VG	Once	.34	.32
		Twice	.45	.34
		Total	.40	.33
	Total		.26	.24
Neutral	V	Once	.11	.18
		Twice	.14	.18
		Total	.12	.18
	G	Once	.22	.20
		Twice	.15	.18
		Total	.19	.19
	VG	Once	.34	.30
		Twice	.37	.31
		Total	.35	.30
	Total		.22	.22

Positive	V	Once	.06	.13
		Twice	.10	.16
		Total	.08	.14
	G	Once	.08	.12
		Twice	.11	.14
		Total	.10	.13
	VG	Once	.21	.19
		Twice	.27	.27
		Total	.24	.23
	Total		.14	.17

CDs – Immediate Test

		M	*SD*
V	Negative	.33	.34
	Neutral	.40	.34
	Positive	.45	.34
	Total	.41	.34
G	Negative	.20	.25
	Neutral	.29	.32
	Positive	.20	.23
	Total	.24	.28
VG	Negative	.62	.33
	Neutral	.56	.39
	Positive	.46	.32
	Total	.55	.36
	Negative	.39	.29
Total	Neutral	.42	.34
	Positive	.37	.29

CDs – Delayed Test

Valence	Question	Times Tested	*M*	*SD*
Negative	V	Once	.21	.23
		Twice	.24	.23
		Total	.22	.23
	G	Once	.23	.27
		Twice	.21	.21
		Total	.22	.24
	VG	Once	.39	.35
		Twice	.53	.33
		Total	.46	.34
	Total		.30	.27
Neutral	V	Once	.12	.16
		Twice	.27	.26
		Total	.19	.21
	G	Once	.21	.25
		Twice	.16	.23
		Total	.19	.24
	VG	Once	.37	.31
		Twice	.50	.33
		Total	.43	.32
	Total		.27	.26
Positive	V	Once	.10	.21
		Twice	.16	.21
		Total	.13	.21
	G	Once	.21	.22
		Twice	.14	.19
		Total	.18	.20
	VG	Once	.32	.27

	Twice	.41	.28
	Total	.37	.27
Total		.23	.23

Appendix B

Experiment 2 Acceptance Probabilities

URDs – Immediate Test

		M	*SD*
	Negative	.26	.33
V	Neutral	.39	.36
	Positive	.29	.32
	Total	.32	.34
	Negative	.45	.29
G	Neutral	.50	.30
	Positive	.39	.31
	Total	.45	.30
	Negative	.22	.28
VG	Neutral	.21	.30
	Positive	.25	.34
	Total	.23	.31
	Negative	.31	.43
Total	Neutral	.37	.39
	Positive	.30	.40

URDs – Delayed Test

	Valence	*M*	*SD*	*M*	*SD*	*M*	*SD*
		Tested		Not Tested		Total	
V	Negative	.46	.33	.34	.28	.38	.31
	Neutral	.45	.35	.41	.35	.47	.35
	Positive	.44	.31	.30	.30	.38	.30

	Total	.45	.33	.36	.31	.41	.32
G	Negative	.37	.33	.29	.31	.31	.32
	Neutral	.42	.35	.37	.37	.43	.36
	Positive	.40	.33	.34	.35	.36	.34
	Total	.40	.33	.32	.35	.37	.34
VG	Negative	.30	.26	.29	.30	.29	.28
	Neutral	.46	.32	.50	.40	.49	.36
	Positive	.47	.37	.29	.34	.38	.38
	Total	.39	.33	.37	.36	.38	.35
Total	Negative	.40	.31	.33	.29	.37	.30
	Neutral	.44	.34	.40	.37	.42	.36
	Positive	.42	.33	.35	.33	.37	.33

Targets – Immediate Test

		M	*SD*	*M*	*SD*	*M*	*SD*
		Short Lists		Long Lists		Total	
	Negative	.34	.28	.37	.33	.36	.30
V	Neutral	.30	.30	.26	.33	.28	.31
	Positive	.34	.28	.31	.34	.33	.30
	Total	.33	.29	.31	.33	.32	.30
	Negative	.15	.22	.15	.24	.15	.23
G	Neutral	.17	.25	.11	.18	.14	.21
	Positive	.22	.28	.24	.29	.23	.28
	Total	.18	.25	.17	.24	.17	.25
	Negative	.44	.32	.46	.33	.46	.32
VG	Neutral	.44	.31	.51	.35	.47	.33
	Positive	.50	.37	.51	.39	.50	.38
	Total	.46	.33	.50	.36	.48	.34
	Negative	.31	.27	.33	.30	.32	.29
Total	Neutral	.30	.29	.29	.29	.30	.29

	Positive	.35	.31	.36	.34	.35	.33

Targets – Delayed Test

		M	*SD*	*M*	*SD*	*M*	*SD*
		Not Previously Tested					
		Short Lists		Long Lists		Total	
V	Negative	.12	.19	.18	.29	.15	.24
	Neutral	.11	.17	.11	.15	.11	.16
	Positive	.14	.20	.15	.21	.15	.20
	Total	.12	.18	.14	.22	.13	.20
G	Negative	.48	.29	.19	.24	.33	.27
	Neutral	.42	.30	.18	.26	.30	.28
	Positive	.40	.30	.24	.26	.32	.28
	Total	.43	.30	.20	.25	.32	.28
VG	Negative	.41	.30	.33	.32	.37	.31
	Neutral	.31	.31	.19	.23	.25	.28
	Positive	.38	.37	.30	.36	.34	.37
	Total	.37	.32	.27	.31	.32	.32
Total	Negative	.34	.26	.23	.28	.29	.27
	Neutral	.28	.26	.16	.21	.22	.24
	Positive	.31	.29	.23	.28	.27	.29
		Previously Tested					
		Short Lists		Long Lists		Total	
V	Negative	.27	.25	.20	.24	.24	.25
	Neutral	.25	.27	.22	.23	.24	.25
	Positive	.26	.24	.29	.27	.27	.26
	Total	.26	.25	.24	.25	.25	.25
G	Negative	.23	.27	.23	.27	.23	.27
	Neutral	.18	.23	.16	.17	.17	.20
	Positive	.15	.21	.22	.26	.19	.24

	Total	.18	.24	.20	.24	.20	.24
	Negative	.41	.31	.45	.29	.43	.30
VG	Neutral	.24	.25	.30	.29	.27	.27
	Positive	.30	.33	.35	.30	.33	.32
	Total	.31	.30	.37	.30	.34	.30
	Negative	.31	.28	.29	.27	.30	.28
Total	Neutral	.22	.25	.23	.23	.22	.24
	Positive	.24	.26	.29	.28	.26	.27
				Across Prior Testing			
		Short Lists		Long Lists		Total	
	Negative	.20	.22	.19	.27	.19	.25
V	Neutral	.18	.23	.16	.19	.17	.21
	Positive	.20	.22	.22	.24	.21	.23
	Total	.19	.22	.19	.24	.19	.23
	Negative	.36	.28	.21	.26	.28	.27
G	Neutral	.30	.27	.17	.23	.23	.25
	Positive	.28	.26	.23	.26	.25	.26
	Total	.31	.27	.20	.25	.26	.26
	Negative	.41	.31	.39	.31	.40	.31
VG	Neutral	.28	.28	.25	.26	.26	.27
	Positive	.34	.35	.33	.33	.33	.34
	Total	.34	.31	.32	.30	.33	.31
	Negative	.32	.27	.26	.28	.29	.28
Total	Neutral	.25	.26	.19	.23	.22	.25
	Positive	.27	.28	.26	.28	.27	.28

CDs – Immediate Test

		M	*SD*	*M*	*SD*	*M*	*SD*
		Short Lists		Long Lists		Total	
	Negative	.25	.40	.35	.37	.29	.39

V	Neutral	.32	.39	.34	.36	.33	.38
	Positive	.39	.44	.42	.33	.40	.39
	Total	.32	.41	.37	.35	.34	.39
	Negative	.26	.38	.17	.30	.21	.34
G	Neutral	.36	.42	.20	.32	.28	.37
	Positive	.44	.45	.32	.37	.38	.41
	Total	.35	.42	.23	.34	.29	.37
	Negative	.57	.38	.68	.39	.62	.39
VG	Neutral	.55	.39	.57	.43	.56	.41
	Positive	.50	.38	.68	.40	.59	.39
	Total	.54	.38	.64	.40	.59	.40
	Negative	.36	.39	.41	.36	.38	.38
Total	Neutral	.41	.40	.40	.37	.40	.38
	Positive	.44	.40	.43	.37	.43	.38

CDs – Delayed Test

		M	*SD*	*M*	*SD*	*M*	*SD*
				Not Previously Tested			
		Short Lists		Long Lists		Total	
	Negative	.14	.17	.29	.32	.21	.25
V	Neutral	.14	.20	.36	.33	.33	.27
	Positive	.20	.26	.41	.31	.36	.29
	Total	.16	.22	.35	.32	.26	.27
	Negative	.24	.26	.43	.42	.25	.26
G	Neutral	.19	.28	.40	.39	.30	.34
	Positive	.26	.29	.37	.38	.25	.34
	Total	.23	.28	.40	.39	.32	.34
	Negative	.35	.37	.37	.37	.31	.37
VG	Neutral	.26	.28	.24	.31	.32	.30
	Positive	.31	.36	.34	.32	.32	.34

	Total	.30	.34	.31	.33	.31	.34
	Negative	.24	.27	.36	.37	.30	.33
Total	Neutral	.20	.25	.33	.34	.27	.26
	Positive	.26	.30	.37	.34	.32	.32
		Previously Tested					
		Short Lists		Long Lists		Total	
	Negative	.29	.32	.30	.33	.29	.33
V	Neutral	.32	.32	.28	.29	.36	.31
	Positive	.37	.33	.30	.25	.53	.29
	Total	.33	.32	.29	.29	.31	.31
	Negative	.45	.38	.27	.31	.30	.35
G	Neutral	.21	.33	.22	.25	.22	.29
	Positive	.41	.41	.27	.32	.36	.37
	Total	.36	.39	.25	.29	.31	.34
	Negative	.48	.34	.58	.35	.34	.35
VG	Neutral	.36	.32	.36	.34	.34	.33
	Positive	.42	.33	.42	.38	.42	.35
	Total	.42	.33	.45	.37	.44	.35
	Negative	.41	.35	.38	.33	.39	.34
Total	Neutral	.30	.32	.28	.29	.29	.32
	Positive	.40	.36	.33	.32	.37	.34
		Across Prior Testing					
		Short Lists		Long Lists		Total	
	Negative	.21	.25	.29	.33	.25	.29
V	Neutral	.23	.26	.32	.31	.28	.29
	Positive	.29	.30	.36	.28	.32	.29
	Total	.24	.27	.32	.31	.28	.29
	Negative	.35	.32	.35	.37	.35	.35
G	Neutral	.20	.31	.31	.32	.26	.32
	Positive	.24	.35	.32	.35	.33	.35
	Total	.30	.34	.33	.34	.31	.34

	Negative	.41	.36	.48	.36	.45	.36
VG	Neutral	.31	.30	.30	.33	.30	.30
	Positive	.36	.35	.38	.35	.37	.36
	Total	.36	.34	.38	.35	.37	.35
	Negative	.32	.34	.37	.35	.35	.35
Total	Neutral	.25	.29	.31	.32	.28	.31
	Positive	.33	.33	.35	.33	.34	.33

Appendix C

Experiment 3 Acceptance Probabilities

URDs – Immediate Test

		M	*SD*
	Negative	.38	.24
V	Neutral	.31	.30
	Positive	.24	.30
	Total	.30	.31
	Negative	.41	.33
G	Neutral	.35	.29
	Positive	.41	.29
	Total	.39	.30
	Negative	.22	.29
VG	Neutral	.27	.34
	Positive	.33	.32
	Total	.28	.32
	Negative	.34	.29
Total	Neutral	.31	.32
	Positive	.32	.30

URDs – Delayed Test

	Valence	*M*	*SD*	*M*	*SD*	*M*	*SD*
		Tested		Not Tested		Total	
V	Negative	.41	.33	.29	.28	.35	.31
	Neutral	.45	.35	.41	.35	.42	.35
	Positive	.49	.31	.30	.30	.41	.31
	Total	.45	.33	.33	.31	.39	.32
G	Negative	.37	.33	.29	.31	.33	,32
	Neutral	.47	.35	.37	.37	.42	.36
	Positive	.40	.33	.34	.35	.37	.34
	Total	.41	.33	.32	.35	.37	.34
VG	Negative	.30	.26	.25	.30	.38	.38
	Neutral	.46	.32	.50	.40	.31	.36
	Positive	.43	.37	.29	.34	.35	.36
	Total	.39	.33	.35	.36	.37	.35
Total	Negative	.35	.32	.27	.30	.32	.31
	Neutral	.41	.35	.41	.37	.41	.36
	Positive	.42	.34	.33	.34	.37	.34

Targets – Immediate Test

		M	*SD*	*M*	*SD*	*M*	*SD*
		Short Lists		Long Lists		Total	
	Negative	.26	.26	.28	.29	.28	.27
V	Neutral	.25	.28	.33	.31	.29	.30
	Positive	.42	.30	.33	.33	.38	.32
	Total	.30	.28	.32	.31	.32	.30
	Negative	.24	.28	.23	.32	.23	.30
G	Neutral	.17	.25	.22	.28	.20	.27
	Positive	.13	.18	.15	.24	.14	.21
	Total	.18	.25	.20	.29	.19	.27
	Negative	.48	.35	.54	.35	.51	.35
VG	Neutral	.38	.35	.47	.36	.42	.36
	Positive	.41	.40	.40	.37	.40	.39
	Total	.42	.36	.47	.36	.45	.36
	Negative	.33	.30	.35	.32	.34	,31
Total	Neutral	.27	.29	.34	.32	.30	.31
	Positive	.32	.29	.29	.31	.31	.30

Targets – Delayed Test

		M	*SD*	*M*	*SD*	*M*	*SD*
				Not Previously Tested			
		Short Lists		Long Lists		Total	
	Negative	.15	.20	.22	.26	.19	.23
V	Neutral	.08	.15	.19	.26	.14	.15
	Positive	.16	.19	.24	.25	.21	.20
	Total	.13	.18	.21	.25	.17	.18
	Negative	.43	.35	.18	.22	.31	.29
G	Neutral	.42	.28	.20	.26	.31	.27
	Positive	.49	.29	.19	.22	.34	.32

	Total	.45	.31	.19	.23	.32	.32
	Negative	.34	.33	.26	.28	.30	.32
VG	Neutral	.31	.32	.22	.26	.26	.29
	Positive	.36	.32	.29	.34	.33	.35
	Total	.34	.32	.26	.29	.30	.31
	Negative	.31.	.29	.22	.25	.27	.27
Total	Neutral	.27	.25	.20	.26	.24	.26
	Positive	.34	.27	.24	.27	.29	.27
		Previously Tested					
		Short Lists		Long Lists		Total	
	Negative	.32	.27	.31	.28	.32	.28
V	Neutral	.30	.26	.30	.31	.30	.29
	Positive	.36	.21	.27	.25	.32	.23
	Total	.32	.25	.29	.28	.31	.27
	Negative	.22	.27	.21	.27	.22	.27
G	Neutral	.22	.27	.21	.24	.22	.26
	Positive	.20	.23	.23	.26	.21	.25
	Total	.21	.26	.22	.26	.22	.26
	Negative	.35	.34	.36	.29	.36	.32
VG	Neutral	.34	.34	.34	.29	.34	.32
	Positive	.36	.33	.38	.36	.37	.35
	Total	.35	.34	.36	.31	.36	.32
	Negative	.30	.29	.29	.28	.30	.29
Total	Neutral	.29	.29	.28	.28	.28	.29
	Positive	.31	.26	.29	.29	.30	.27
		Across Prior Testing					
		Short Lists		Long Lists		Total	
	Negative	.24	.24	.27	.27	.25	.25
V	Neutral	.19	.21	.25	.29	.22	.22
	Positive	.26	.20	.25	.25	.26	.23
	Total	.23	.22	.25	.27	.24	.24

	Negative	.33	.31	.20	.25	.26	.29
G	Neutral	.32	.28	.21	.25	.26	.27
	Positive	.35	.26	.21	.24	.28	.27
	Total	.33	.29	.20	.25	.27	.28
	Negative	.35	.34	.32	.29	.33	.34
VG	Neutral	.32	.33	.28	.28	.30	.32
	Positive	.36	.33	.33	.35	.35	.34
	Total	.34	.33	.31	.30	.33	.33
	Negative	.30	.29	.26	.27	.28	.39
Total	Neutral	.28	.27	.24	.27	.26	.37
	Positive	.32	.27	.27	.28	.29	.28

CDs – Immediate Test

		M	*SD*	*M*	*SD*	*M*	*SD*
		Short Lists		Long Lists		Total	
	Negative	.47	.43	.66	.36	.71	.40
V	Neutral	.29	.51	.68	.34	.57	.43
	Positive	.62	.42	.54	.31	.66	.37
	Total	.65	.48	.63	.34	.65	.41
	Negative	.34	.40	.25	.39	.29	.40
G	Neutral	.26	.38	.19	.34	.22	.36
	Positive	.31	.37	.24	.32	.28	.35
	Total	.30	.38	.22	.35	.27	.25
	Negative	.63	.40	.54	.45	.58	.43
VG	Neutral	.51	.38	.63	.41	.57	.40
	Positive	.44	.41	.56	.37	.50	.39
	Total	.53	.40	.58	.41	.55	.41
	Negative	.58	.41	.48	.40	.53	.41
Total	Neutral	.41	.42	.50	.36	.45	.39
	Positive	.51	.40	.44	.33	.48	.37

CDs – Delayed Test

		M	*SD*	*M*	*SD*	*M*	*SD*
		Not Previously Tested					
		Short Lists		Long Lists		Total	
	Negative	.18	.24	.32	.35	.25	.30
V	Neutral	.09	.18	.27	.34	.19	26
	Positive	.29	.33	.36	.37	.33	.35
	Total	.18	.26	.32	.35	.26	.31
	Negative	.25	.29	.29	.36	.27	.33
G	Neutral	.19	.24	.35	.38	.27	.31
	Positive	.26	.29	.33	.38	.29	.34
	Total	.23	.27	.32	.37	.28	.32
	Negative	.30	.34	.35	.30	.32	.32
VG	Neutral	.29	.35	.34	.38	.31	.37
	Positive	.46	.39	.46	.40	.46	.40
	Total	.34	.36	.38	.36	.37	.36
	Negative	.24	.29	.32	.34	.28	.32
Total	Neutral	.19	.26	.32	.37	.26	.32
	Positive	.34	.34	.38	.38	.36	.36
		Previously Tested					
		Short Lists		Long Lists		Total	
	Negative	.46	.32	.29	.32	.38	.32
V	Neutral	.39	.36	.32	.29	.35	.33
	Positive	.41	.34	.28	.26	.35	.30
	Total	.42	.34	.30	.29	.36	.32
	Negative	.29	.35	.24	.27	.26	.31
G	Neutral	.32	.36	.24	.31	.28	.34
	Positive	.30	.37	.24	.31	.27	.34
	Total	.31	.35	.24	.29	.27	.32

	Negative	.45	.34	.47	.37	.46	.36
VG	Neutral	.40	.36	.52	.37	.46	.37
	Positive	.41	.33	.56	.37	.49	.35
	Total	.42	.34	.51	.37	.47	.36
	Negative	.40	.34	.33	.32	.37	.33
Total	Neutral	.37	.36	.36	.32	.37	.34
	Positive	.37	.35	.36	.31	.38	.33
				Across Prior Testing			
		Short Lists		Long Lists		Total	
	Negative	.32	.28	.31	.34	.31	.30
V	Neutral	.24	.27	.30	.32	.27	.30
	Positive	.35	.34	.32	.32	.34	.33
	Total	.31	.30	.31	.32	.31	.31
	Negative	.27	.32	.26	.32	.27	.32
G	Neutral	.26	.31	.30	.34	.28	.33
	Positive	.28	.33	.28	.34	.28	.34
	Total	.27	.32	.28	.33	.27	.33
	Negative	.37	.34	.41	.34	.39	.34
VG	Neutral	.35	.36	.43	.38	.39	.37
	Positive	.44	.36	.51	.39	.47	.38
	Total	.39	.35	.45	.37	.42	.37
	Negative	.32	.32	.33	.33	.32	.33
Total	Neutral	.28	.31	.34	.35	.31	.33
	Positive	.36	.35	.37	.35	.36	.35

Appendix D

Experiment 4 Acceptance Probabilities

URDs – Immediate Test

	Arousal	*M*	*SD*
V	Low	.31	.33
	Medium	.31	.33
	High	.31	.26
	Total	.31	.31
G	Low	.43	.34
	Medium	.36	.35
	High	.41	.38
	Total	.39	.35
VG	Low	.26	.32
	Medium	.24	.27
	High	.40	.43
	Total	.27	.32
	Low	.33	.33
Total	Medium	.30	.31
	High	.37	.36

URDs – Delayed Test

	Arousal	*M*	*SD*	*M*	*SD*	*M*	*SD*
		Previously Tested		Not Previously Tested		Total	
V	Low	.41	.35	.30	.29	.36	.33
	Medium	.45	.32	.35	.31	.40	.32
	High	.52	.30	.31	.23	.42	.26
	Total	.45	.32	.33	.29	.38	.31
G	Low	.44	.29	.31	.35	.38	.33

	Medium	.38	.35	.35	.35	.35	.35
	High	.42	.32	.31	.31	.42	.32
	Total	.41	.33	.33	.34	.37	.34
VG	Low	.38	.37	.24	.33	.31	.36
	Medium	.42	.34	.42	.37	.42	.36
	High	.41	.38	.33	.34	.40	.37
	Total	.40	.35	.35	.36	.37	.36
Total	Low	.41	.35	.29	.33	.35	.34
	Medium	.44	.34	.36	.34	.40	.34
	High	.45	.34	.34	.28	.40	.33

Targets – Immediate Test

	Arousal	*M*	*SD*	*M*	*SD*	*M*	*SD*
		Short		Long		Total	
V	Low	.30	.29	.32	.34	.31	.31
	Medium	.31	.28	.32	.30	.31	.29
	High	.29	.30	.29	.27	.29	.29
	Total	.30	.28	.32	.31	.30	.30
G	Low	.15	.22	.21	.28	.18	.25
	Medium	.20	.26	.23	.31	.22	.28
	High	.20	.24	.10	.19	.15	.22
	Total	.18	.25	.20	.29	.19	.27
VG	Low	.40	.38	.46	.38	.43	.38
	Medium	.44	.35	.49	.33	.47	.34
	High	.40	.41	.46	.42	.43	.41
	Total	.42	.36	.47	.36	.44	.36
	Low	.27	.30	.27	.34	.27	.32
Total	Medium	.31	.32	.25	.31	.28	.31
	High	.33	.34	.25	.29	.29	.32

Targets – Delayed Test

		M	*SD*	*M*	*SD*	*M*	*SD*
				Not Previously Tested			
		Short Lists		Long Lists		Total	
V	Low	.11	.17	.22	.26	.16	.23
	Medium	.11	.17	.21	.25	.16	.22
	High	.24	.20	.24	.24	.24	.22
	Total	.13	.18	.21	.25	.19	.22
G	Low	.34	.30	.24	.27	.30	.29
	Medium	.47	.30	.16	.20	.32	.25
	High	.55	.32	.19	.26	.37	.39
	Total	.45	.31	.19	.23	.33	.32
VG	Low	.36	.37	.29	.34	.33	.36
	Medium	.34	.31	.23	.25	.29	.28
	High	.27	.25	.27	.30	.27	.28
	Total	.34	.32	.26	.29	.29	.32
Total	Low	.27	.28	.25	.28	.26	.28
	Medium	.31	.26	.20	.23	.26	.25
	High	.35	.26	.23	.27	.29	.27
				Previously Tested			
		Short Lists		Long Lists		Total	
V	Low	.29	.24	.29	.26	.29	.25
	Medium	.33	.25	.30	.29	.32	.27
	High	.38	.25	.28	.28	.33	.27
	Total	.32	.25	.29	.28	.31	.26
G	Low	.14	.18	.20	.26	.17	.22
	Medium	.28	.29	.24	.26	.26	.27
	High	.16	.19	.15	.25	.15	.22
	Total	.21	.26	.22	.26	.19	.23
	Low	.41	.37	.40	.34	.40	.26

VG	Medium	.31	.30	.34	.30	.32	.30
	High	.38	.37	.34	.31	.36	.34
	Total	.35	.34	.36	.31	.36	.30
	Low	.28	.26	.30	.29	.29	.28
Total	Medium	.30	.28	.29	.28	.20	.28
	High	.30	.27	.26	.28	.28	.28

		Across Prior Testing					
		Short Lists		Long Lists		Total	
	Low	.20	.21	.25	.26	.23	.24
V	Medium	.22	.22	.25	.27	.24	.25
	High	.31	.23	.26	.26	.28	.25
	Total	.24	.22	.26	.26	.25	.25
	Low	.24	.24	.22	.27	.23	.26
G	Medium	.37	.30	.20	.23	.29	.27
	High	.35	.26	.17	.26	.26	.26
	Total	32	.26	.20	.25	.26	.26
	Low	.38	.37	.34	.34	.36	.36
VG	Medium	.32	.31	.28	.28	.30	.30
	High	.32	.31	.31	.31	.32	.31
	Total	.34	.33	.31	..30	.33	.32
	Low	.27	.27	.27	.29	.27	.28
Total	Medium	.31	.27	.25	.26	.28	.27
	High	.33	.27	.25	.28	.29	.28

CDs – Immediate Test

	Arousal	*M*	*SD*	*M*	*SD*	*M*	*SD*
		Short		Long		Total	
V	Low	.36	.44	.35	.34	.37	.39
	Medium	.40	.41	.40	.34	.43	.38
	High	.48	.36	.36	.37	.42	.37

	Total	.44	.41	.38	.34	.41	.38
G	Low	.37	.41	.18	.31	.28	.36
	Medium	.25	.36	.24	.39	.25	.37
	High	.34	.40	.25	.32	.30	.37
	Total	.30	.38	.22	.35	.27	.37
VG	Low	.50	.41	.57	.43	.53	.42
	Medium	.59	.37	.63	.39	.61	.38
	High	.41	.46	.42	.43	.41	.45
	Total	.53	.40	.58	.41	.56	.41
	Low	.42	.42	.37	.36	.39	.39
Total	Medium	.43	.38	.42	.37	.43	.38
	High	.41	.40	.34	.37	.38	.39

CDs – Delayed Test

		M	*SD*	*M*	*SD*	*M*	*SD*
				Not Previously Tested			
		Short Lists		Long Lists		Total	
	Low	.15	.25	.28	.36	.21	.31
V	Medium	.18	.27	.35	.36	.27	.32
	High	.22	.28	.27	.30	.25	.29
	Total	.18	.26	.32	.35	.24	.31
	Low	.18	.23	.27	.36	.22	.30
G	Medium	.26	.30	.38	.37	.32	.34
	High	.22	.23	.25	.35	.24	.29
	Total	.23	.27	.32	.37	.26	.32
	Low	.39	.41	.41	.44	.40	.43
VG	Medium	.30	.34	.35	.31	.32	.33
	High	.38	.36	.40	.37	.39	.37
	Total	.34	.36	.38	.36	.37	.36
	Low	.24	.30	.32	.39	.28	.35

Total	Medium	.25	.30	.36	.35	.30	.33
	High	.27	.29	.31	.34	.29	.32
		Previously Tested					
		Short Lists		Long Lists		Total	
	Low	.39	.36	.35	.31	.37	.34
V	Medium	.45	.34	.26	.28	.36	.31
	High	.36	.32	.32	.27	.34	.30
	Total	.42	.34	.30	.29	.36	.32
	Low	.35	.37	.25	.32	.30	.35
G	Medium	.31	.36	.25	.28	.28	.32
	High	.20	.29	.21	.30	.20	.30
	Total	.31	.35	.24	.29	.26	.32
	Low	.44	.39	.56	.40	.50	.40
VG	Medium	.42	.32	.50	.34	.46	.33
	High	.36	.33	.46	.38	.41	.36
	Total	.42	.34	.51	.37	.46	.36
	Low	.40	.38	.39	.36	.39	.37
Total	Medium	.40	.34	.34	.30	.37	.32
	High	.31	.32	.33	.33	.32	.33
		Across Prior Testing					
		Short Lists		Long Lists		Total	
	Low	.27	.31	.31	.34	.29	.33
V	Medium	.32	.30	.31	.32	.31	.31
	High	.29	.30	.30	.29	.29	.30
	Total	.29	.30	.31	.32	.30	.31
	Low	.27	.30	.26	.34	.26	.32
G	Medium	.28	.34	.31	.33	.30	.34
	High	.21	.26	.23	.33	.22	.29
	Total	.25	.30	.27	.33	.26	.32
	Low	.42	.40	.49	.42	.45	.41
VG	Medium	.36	.33	.43	.33	.39	.33

	High	.37	.35	.43	.38	.40	.37
	Total	.38	.35	.45	.37	.42	.36
	Low	.32	.34	.35	.38	.34	.36
Total	Medium	.32	.32	.35	.33	.34	.33
	High	.29	.31	.31	.34	.30	.33