I BELIEVE IN A THING CALLED EMOTION: EFFECTS OF VALENCE ON MEMORY MONITORING AND PERFORMANCE

by © Landon A. Churchill

A Thesis submitted to the School of Graduate Studies

in partial fulfillment of the requirements

for the degree of

Masters of Science in Experimental Psychology

Department of Psychology

Faculty of Science

Memorial University of Newfoundland

October 2017

St. John's   Newfoundland and Labrador

Abstract

It is generally well known that emotion can benefit memory; however, much less is known about how emotion influences metamemory. The current thesis examined how list composition and emotion influences individuals' metamnemonic judgments and performance in a free recall task. Participants studied lists of words that varied in emotional valence and made immediate judgments of learning (JOLs) after each word. Valence was manipulated in a mixed-list design in Experiment 1, and a pure-list design in Experiment 2, while arousal was held constant. It was expected that valence would affect participants' JOLs and recall performance, but only in Experiment 1. Consistent with this, emotional words were given higher JOLs and were recalled better than neutral words in Experiment 1, while no such differences were observed in Experiment 2. Results suggest that the metamnemonic effect of emotion is limited to mixed-list designs, and likely depends on participants' beliefs about how emotion influences memory.

Keywords: metamemory, JOLs, recall, emotion

Acknowledgements

I would like to express the most sincere gratitude to my supervisor, Dr. Hourihan, for being a remarkable mentor during the past few years. Without her wisdom and support none of this would have been possible. I would also like to thank my supervisory committee for assisting me with the writing process; Brandon Slaney, for troubleshooting during programming; and Whitney Wilcott-Benoit, for coding the questionnaire responses. Additional thanks are extended to NSERC, which provided the funding necessary to finish this project.

I would like to dedicate this work to the memory of my late grandfather, who firmly believed “without knowledge, we have nothing”.

Co-authorship Statement:

This project was initiated as a follow-up to a previous study conducted by my supervisor. I was responsible for programming both experiments using E-Prime software, and collecting, coding, and analysing the data using SPSS and Microsoft Excel. This manuscript is an original work, with suggestions from my supervisory committee incorporated.

Table of Contents

List of Tables

List of Figures

List of Appendices

I Believe in a Thing Called Emotion:

The Influence of Valence on Memory Monitoring and Performance

The influence of emotion on memory has long been a topic of interest to psychologists and laypersons alike. Notably, in an early essay titled "What is an Emotion?" William James (1884, p. 300) asked, "how much of our mental life is knit up with our corporeal frame?". Questions such as these suggest that emotion plays an interactive role with respect to both our physiological and cognitive processes. Indeed, early investigations of emotion's influence on cognitive mechanisms focused on whether events that are encoded in a certain state are easier to retrieve in the same state or in a different state (e.g., Network Theory of Emotions, Bower, 1981; Affect Infusion Model, Forgas, 1995; Mood-Dependent Memory, McGeoch, 1942). We know that memory is better for items tested in the same context as they are encoded (Context-Dependent Learning, Eich, 1980; Encoding Specificity Principle, Tulving & Thomson, 1973), thus the notion that theories of emotion have followed this premise is not surprising.

Investigations of the influence of emotion on memory also involve memory for emotional information while in a neutral state. Such research has primarily manipulated the emotional content of stimuli such as words (e.g., Hourihan, Fraundorf, & Benjamin, 2017; Kensinger & Corkin, 2003; Talmi & Moscovitch, 2004) and pictures (e.g., Bradley, Greenwald, Petry, & Lang, 1992; Hourihan & Bursey, 2016; Talmi & McGarry, 2012), and generally converge on the finding that emotional stimuli are better remembered than non-emotional stimuli (for a review, see Kensinger, 2009). However, inconsistencies have arisen with respect to which *dimension* of emotion (i.e., arousal or valence) primarily influences memory. As such, one purpose of the current set of experiments was to further investigate this ambiguity.

Recently, researchers have been interested in the influence of emotion on metacognition (i.e., monitoring and control of cognition; e.g., Efklides & Petkaki, 2005; Hourihan et al., 2017; Koriat & Nussinson, 2009; Nomi, Rhodes, & Cleary, 2013; Tauber & Dunlosky, 2012; Zimmerman & Kelley, 2010). As aforementioned, the effect of arousal and valence on memory is, in reality, quite inconsistent; however, conclusions drawn from the effect of emotion on metamemory are generally unvarying. Specifically, participants give higher predictions of future recall for emotional stimuli relative to neutral stimuli, even when their memory performance is unaffected by emotionality (Hourihan et al., 2017; Nomi et al., 2013). Critically, however, the mechanisms underlying this effect remain unclear. It has been speculated (e.g., Hourihan et al., 2017), but remains debatable as to whether the effects of emotion on metamemory are primarily influenced directly via intrinsic qualities of emotional stimuli (e.g., imageability), or if metacognitive knowledge (e.g., beliefs about how memory works) acts as a mediating factor. Thus, the primary objective of this thesis was to further investigate the proposed mechanisms responsible for the metamnemonic effect of emotional stimuli.

In the rest of the introduction I will discuss the aspects that define emotional stimuli and their role in metamemory processes. Then I will discuss the role of list structure in making judgments of learning. Finally, I describe two experiments that bring these literatures together, and propose that the effects of emotion on metamemory are primarily driven by list structure, coupled with participants' beliefs regarding the differential memorability of emotional and non-emotional information.

1.1 Scope of emotionality: Arousal and valence

The examination of emotion begins with defining two aspects of emotional stimuli: arousal and valence. Word stimuli with these qualities are commonly quantified by numbers from 1 (low) to 8 (high), based on a large sample norming study (Affective Norms for English Words [ANEW], Bradley & Lang, 2010). The dimension of arousal indicates the level of 'energy' associated with an emotional word, from low to high; valence represents the 'pleasantness' or 'attractiveness' of a word, from negative to positive. For example, "excitement" is a high arousal, positive word; "bored" is a low arousal, negative word. It has been suggested that emotional information is more accurately retained (compared to non-emotional information) primarily due to the dimension of arousal (Kensinger, 2004; MacLeod & Matthews, 2004). More specifically, research has shown that arousing information is more likely to be attended to, facilitating encoding of said information (Kensinger, 2009; MacLeod & Matthews, 2004). Once encoded, arousing information may also be more likely to be retrieved, though evidence fails to explain specifically *how* emotion influences retrieval (Hamann, 2001). It also has been postulated that arousing information is remembered more accurately than neutral information because it engages different metacognitive processes than the mechanisms responsible for the memorability of non-emotional information (e.g., additional rehearsal, Reisberg & Heuer, 2004). Specifically, participants may have more *intent* to focus their attention during encoding when information is more arousing (Efklides, 2016).

Although this argument exists, research also indicates that arousing information does not provide any benefit to the *degree* of detail encoded (Mather, 2007). Instead, effects of arousal on memory may be best characterized by focal enhancements (Mather, 2007; Reisberg & Heuer, 2004). For example, participants with high anxiety have been shown to focus more on negative information compared to individuals expressing low levels of anxiety, resulting in better memory for this information (e.g., Ferguson, Moghaddam, & Bibby, 2007; MacLeod & Mathews, 2004). Although these postulated theories are moderately independent from one another, they all converge on the notion that some aspects of an emotional experience are better remembered primarily due to the dimension of arousal.

A closer examination of the effect of valence on recall suggests that it is able to influence memory independently from arousal (e.g., Hourihan et al., 2017; Kensinger & Corkin, 2003). However, discrepant results explaining this dimension of emotion also exist. Similar to the proposed effects of arousal, *negative* valence has been related to focal memory enhancements, whereas *positive* valence has been negatively related to overall memory performance (Kensinger, O'Brien, Swanberg, Garoff-Eaton, & Schacter, 2007). Additionally, neuroimaging studies have found that positively and negatively valenced information activates different regions of the brain (e.g., the amygdala, cingulate gyrus, and bilateral frontal and parietal areas). More specifically, memory for negatively valenced information has been associated with sensory processing areas (compared to positive and neutral valence), whereas memory for positive information has been linked with areas that are active during self-referential processing (Mickley & Kensinger, 2008). Results such as these therefore suggest that the emotional influence on memory may be a product of the varying cognitive processes utilized during initial encoding.

1.2 Emotional influences on metamemory

Metamemory can be simply understood as an individual's introspective knowledge about one's own memory capabilities. Metamemory research focuses on a core set of issues that pertain to people's beliefs about memory, their monitoring of memory, and their control of memory. Flavell (1979) introduced a model of metacognitive monitoring which details four main aspects: metacognitive knowledge, metacognitive experiences, goals, and strategies. Metacognitive knowledge relates to beliefs that individuals may have about their cognitive abilities, and may include beliefs about a task or beliefs regarding the best learning strategies. For example, one may possess a belief that they will perform badly on tests because they believe they are a bad test taker. Metacognitive experiences include contextual cues that arise during a task. Goals and strategies represent metacognitive control processes that interact with the cognitive processes recruited during monitoring. One of the most common measures of memory monitoring is to directly ask participants to rate the likelihood of remembering a stimulus at some point in the future. Such judgments of learning (JOLs) are commonly assessed via responses on an ordinal or ratio scale (Arbuckle & Cuddy, 1969).

As the methods used to evaluate the mechanisms of metamemory mirror those of memory, many of the theories used to explain memory performance are applicable to that of metamemory. One of the most commonly accepted theories of how individuals make JOLs is cue utilization theory (Koriat, 1997). Cue utilization theory states that making a JOL is an inferential process, based on three kinds of cues: intrinsic cues, extrinsic cues, and mnemonic cues. Intrinsic cues are those that are inherent to study items, such as concreteness. Extrinsic cues are those that are associated with the study context itself, such the location of stimuli on a computer screen. Lastly, mnemonic cues are the internal, subjective feelings evoked by items at the time of study, and may rely on a feeling fluency in individuals (i.e., retrieval fluency may be diagnostic of how well items are learned).

Provided the extensive amount of research on the influence of emotional information on memory, it is surprising to see that relatively few studies have attempted to explore the relationship between emotional stimuli and memory monitoring. For example, research has shown that arousal and valence effects are consistent with Koriat's (1997) cue utilization framework such that negative arousal enhances memory for intrinsic features but not extrinsic features (Kensinger et al., 2007). Emotion can be thought of as both an intrinsic and an extrinsic cue: It is an intrinsic feature of word stimuli, but an extrinsic cue in its relationship with mood. Mood-dependent memory effects often mirror those of context-dependent effects, wherein memory performance is optimal if the emotional context (or mood) at test matches the emotional context at encoding; as context is an extrinsic cue, memory for emotional stimuli might also be influenced by "emotion" as a contextual cue. Further, if JOLs are most sensitive to intrinsic cues, they may be influenced by participants' theories regarding study items (e.g., theories regarding the degree of difficulty between items). The influence of mnemonic cues also suggests that JOLs may also be influenced by individual experiences. Given the versatility of emotion as either an intrinsic or extrinsic cue, it is then reasonable to suggest that participants' JOLs may be influenced by both their theories regarding the effects of emotion on memory, and their emotional state during study. In general however, the literature appears to converge on the idea that JOLs are most sensitive to intrinsic cues (e.g., Benjamin, 2003), but not extrinsic cues (e.g., Koriat & Bjork, 1994); therefore, any context-like effect of emotion does not likely influence participants' JOLs.

Recent research by Zimmerman and Kelley (2010) suggests that the experience of reading an emotional word does in fact influence participants' JOLs. The researchers were interested in examining the effect of emotion on memory and memory monitoring, with the premise that emotional words are more easily recalled than non-emotional words. Participants were asked to study a list of neutral, negative, and positive valence word pairs for either a free recall or a cued recall test. As a measure of memory monitoring, participants were asked to provide immediate JOLs after studying each word pair. Zimmerman and Kelley found that participants predicted better recall for emotional word pairs compared to neutral word pairs, and this effect was consistent across test context. Interestingly, memory performance was only influenced by emotion for participants who completed the free recall test. From these results, they concluded that the emotionality of words does influence memory processes; however, the type of contextual test primarily influences performance.

To further extend the results of Zimmerman and Kelley (2010), Tauber and Dunlosky (2012) conducted a study that included both younger and older adults. Following a similar procedure, they replicated the mnemonic benefit of emotional relative to neutral words in free recall; however, they found that only younger adults gave higher JOLs for positive information. This is surprising, as socioemotional selectivity theory would predict the opposite trend. Socioemotional selectivity is a theory of motivation that posits that individuals become more emotionally selective with age; more specifically, selective to positive information compared to neutral information (see Carstensen, 1992). This theory would instead predict that older adults, rather than young adults, should spend less time allocating attention to negative stimuli and show increased recall for positive information. From this, Tauber and Dunlosky proposed that the dimension of arousal might have been responsible for the results, such that the physiological response evoked from reading an emotional word could have potentially influenced participants' responses (e.g., interpreting a word as threatening, and subsequently utilizing the physiological response based on that interpretation as a basis for a higher JOL). Consistent with cue utilization theory (Koriat, 1997), they also suggested that participants may possess certain beliefs about the memorability of emotional stimuli (e.g., a belief that emotional and non-emotional words are remembered differently), and these beliefs may have been driving the results.

Critically, however, many of the conclusions drawn from the aforementioned studies were based on emotional stimuli that not only differed in valence, but also were higher in arousal compared to the neutral items. Based on the confounding of these two dimensions, it remains relatively unclear as to which factor is responsible for establishing the effects of emotion on metamemory.

1.3 Possible explanations: The physiological and cognitive accounts

In order to further delineate the mechanisms responsible for emotion's influence on participants' metacognitive judgments, Hourihan et al. (Experiment 1, 2017) asked participants to study a mixed list of high and low arousal words (equated for valence) for a later recall test. They hypothesized that if the influence of emotion on memory was driven by arousal, high arousal words should be remembered better than low arousal words. Additionally, they reasoned that if the effect of emotion on metamemory is driven by the physiological response produced from reading an arousing word, participants should judge high arousal words to be remembered better than low arousal words, regardless of valence. Surprisingly, the researchers found that arousal did not have an influence on participants' memory performance, but participants predicted that high arousal words would be better recalled than low arousal words. These results provided evidence that arousal can influence participants' metamnemonic judgements. Hourihan et al. (2017) labeled this trend the *physiological account,* such that the experience of reading arousing words would be conductive of a physiological response, and participants could use this response as a cue for future item memorability. Specifically, the influence of arousal is proposed to lead to a feeling of processing fluency in the individual, which in turn would lead to greater predictions of recall, relative to non-arousing stimuli. Support for this account is evident from examinations of neural processes. For example, it has been proposed that the release of stress hormones enhances memory (Kensinger & Corkin, 2003), and also that encoding of arousing information leads to higher activity in the amygdala than that of neutral information (Mickley-Steinmetz & Kensinger, 2004).

Nonetheless, these findings do not fully eliminate the potential influence of factors other than physiological arousal. Because emotionality is an intrinsic component of word stimuli (Koriat, 1997), it is logical to expect that participants may simply interpret high arousal words as being more "emotional" than low arousal words, even without variations in valence. Indeed, Hourihan et al. (2017) reasoned that participants might utilize a belief system about the memorability of emotional words when making predictive judgements.

They postulated that if participants use the intrinsic cue of emotion to make JOLs, then their judgements would be dependent on their interpretation of the memorability of each cue. Their Experiment 2 was conducted to further examine this possibility. Participants studied word lists composed of negative and neutral valence words, equated for arousal. It was predicted that emotional words would be recalled more accurately than neutral words, and participants would report JOLs that corresponded to this difference in emotion. The results yielded an effect of emotion, such that negative valence words were better recalled than neutral valence words. JOLs were also sensitive to this; participants correctly predicted negative valence words to be better recalled than neutral words. Critically, this effect emerged even though the word stimuli were equated on the dimension of arousal, providing evidence against the physiological account. The researchers concluded that participants based their responses on a belief that emotional information should benefit recall, and termed this the *cognitive account.* Ultimately, these results suggest that the effects of emotion on memory and metamemory may instead be determined via a conscious interpretation of the stimuli, rather than physiological responses.

1.4 The influence of list structure

Notably, prior investigations of the benefit to memory for emotional information have been mixed-list designs where different ‘types’ of stimuli are compared in a single context (e.g., Dewhurst & Parry, 2000). We know from other cognitive phenomena (e.g., the Generation Effect, Slamecka & Graf, 1979; the Production Effect, MacLeod, Gopie, Hourihan, Neary, & Ozubko, 2010) that certain memory effects appear to be limited to these designs, as opposed to pure designs. For example, both the Generation and the Production effects state that generated and produced words are remembered better than non-generated and non-produced words, respectively. Critically, these effects are evident when both generated (or produced) items and non-generated (or non-produced) items are studied together (as in mixed-list designs), but fail to be observed in pure-list designs (but see Fawcett, 2013). Additionally, the Perceptual Fluency Hypothesis states that items that are easier to perceive at study are more memorable than items that are perceived as less fluent (e.g., Reber, Winkielman, & Schwarz, 1998). This is also applicable in studies of metamemory: Susser, Mulligan, and Besken (2013) found that participants give higher JOLs to items that they perceive as more fluent. Taken together, these results indicate that JOLs may reflect a comparative process based on participants' perceived fluency between "types" of stimuli encountered.

Following this logic, Hourihan et al. reasoned that their results (Experiment 2, 2017) might have been due to participants' comparing emotional and neutral valence words in a mixed-list design. Specifically, they suggested that negative words might have been seen as distinctive compared to neutral words. In order to more precisely investigate the combinatory effects of list structure and emotion on participants' metacognitive judgements, Hourihan et al. conducted a third experiment using a continuous-levels design. Here, participants studied words that were selected from the entire range of valence and arousal values, from low to high. The premise was that if participants rely on the relative fluency between items, then emotional words must appear relatively distinctive compared to a "background" of neutral words (Schmidt, 1991; Talmi, 2013). Indeed, multi-level modelling yielded no effect of valence or arousal on participants'

JOLs or their memory performance. By using a design that no longer allowed emotional items to be distinctive from neutral items, the effect of emotion on metamemory (and memory) was eradicated. Hourihan et al. concluded that it appeared as though participants assigned higher JOLs to emotional words based on an explicit belief that these stimuli simply should be more memorable. This additionally lends support to their cognitive account, such that the metamnemonic benefit of emotion is proposed to rely upon both a mixed list composition, and participants' cognitive appraisal of the study items.

1.5 Current thesis

Recent evidence appears to converge on the notion that the effects of emotion on metamemory are dependent on the experimental design (e.g., Hourihan et al., 2017; Tauber & Dunlosky, 2012). As aforementioned, Hourihan et al. (2017) found a metamnemonic effect of emotion using negative and neutral valence words in a mixed-list context (Experiment 2), but not with a pure list of words that varied in both arousal and valence (Experiment 3), suggesting that the metamnemonic effect of emotion may indeed be limited (or at least moderated) by the experimental design. If JOLs for emotional words rely upon the relative distinctiveness created from a mixed list of emotional and neutral words, then a further examination of how JOLs vary for mixed and pure lists is required to fully support their proposed cognitive account. For example, it is still unknown as to whether the effect is replicable with respect to positive and neutral lists, while keeping arousal constant.

If the effect of emotion on JOLs is moderated by the experimental design, then participants in mixed-list designs are expected to give higher JOLs to emotional items compared to neutral items, due to the relative distinctiveness of emotional words. Additionally, however, Hourihan et al. (2017) suggested that the metamnemonic effect of emotion might be coupled with participants' explicit belief that emotional information is remembered differently than neutral information. In order to fully assess this proposition, it is essential to explicitly query participants in a post-test questionnaire. It is possible that participants could actually give higher JOLs to emotional (both negative and positive) words, even when studied in a pure list, if they are able to enact their beliefs about the effects of emotion on memory.

The current study compared recall performance and JOLs of emotional and neutral words. Given that valence has been shown to affect performance independently from arousal (e.g., Hourihan et al., 2017; Kensinger, 2009) the dimension of valence was manipulated with arousal equated (at the midpoint). In Experiment 1, each participant studied two mixed lists of words—negative and neutral valence, and positive and neutral valence—and made immediate JOLs after each word. Experiment 2 was procedurally identical to the first, with the exception that each participant studied two lists composed of only one "type" of emotional word— negative, positive, or neutral. At the end of each experiment, participants were asked to complete a paper-and-pencil questionnaire regarding their thoughts and beliefs about emotion and memory, and what they noticed about the words in the experiment (see Appendix B).

Consistent with evidence for the benefit of emotional information to memory (e.g., Hourihan et al., 2017; Kensinger, 2009; Tauber & Dunlosky, 2010), it is expected that participants will recall more emotional words than neutral words. Due to inconsistent conclusions regarding the mnemonic effect of valence (e.g., Kensinger & Corkin, 2003;

Kensinger et al., 2007), it is expected that the “emotionality” of words in general will serve as a cue to participants; therefore, performance in both experiments is not expected to differ based on whether participants studied positive or negative valence words. In terms of participants’ metamnemonic judgements, it is expected that the effect of emotion on participants’ JOLs will depend on list composition, coupled with their explicit noticing of emotional words. That is, when there are mixed-valence lists (as in Experiment 1), JOLs should be higher for emotional words compared to neutral words; conversely, when participants study pure lists (as in Experiment 2), there should be no observable differences in participants’ JOLs. The post-test questionnaire is included to explicitly assess participants’ beliefs and their interpretation of the experimental design. These results are expected to support the cognitive account of the emotional JOL effect; participants should report that they noticed emotional and neutral words in the mixed lists (Experiment 1) and emotional pure lists (Experiment 2), but those who study the neutral pure list (Experiment 2) should not. Importantly, when participants are asked about the general effect of emotion on memory, it is expected that all will report that emotional information is more memorable than non-emotional information.

It may be noted that performance may be related to individual differences in personality or anxiety level. It is important to consider individual differences in affect when examining the influence of emotion on memory and memory monitoring, as the degree to which memory is incremented (or impaired) may vary between participants. Importantly, a central tenant of theories regarding the influence of emotion on memory is that mood is a significant moderator; the relationship between positive information and memory is one that relies upon mood as a mediator of attention and performance. More specifically, positive stimuli have been proposed to invoke a positive mood, which leads individuals to attend to events more broadly and process information heuristically (e.g., Bless, Clore, Schwarz, Golisano, Rabe, & Wolk, 1996; Kensinger et al., 2007; Storbeck & Clore, 2005). As an attempt to measure mood congruency, the post-test questionnaire includes the Positive and Negative Affect Schedule (PANAS, Watson, Clark, & Tellegen, 1988). The PANAS is a clinical measure that requires participants to read a list of 20 words that describe different positive and negative emotions, and indicate how they are feeling at a given time on a scale from 1 ("very slightly or not at all") to 5 ("extremely"). This measure will be used to show whether participants' moods are predictive of their responses. For example, emotional words (either negative or positive) may be more salient to participants who are in a congruent mood, leading these words to be perceived as more memorable (relative to words incongruent with their mood). If scores on the positive and negative subscales of the PANAS were reliability correlated with metamemory or memory performance, this would suggest that participants' performance might be primarily influenced by their mood, rather than the experimental design (e.g., Egidi & Gerrig, 2009; Fieldler & Hutter, 2013).

2.1 Experiment 1

Experiment 1 was conducted to investigate whether the metamnemonic benefit of emotional words is driven by their relative distinctiveness compared to neutral words. A mixed-list design was used; emotional words were selected to vary in valence while arousal was held constant (at the midpoint). Participants studied two mixed lists and were asked to provide immediate JOLs after each word. Following each list, they were asked to complete a free recall test. If the effects of emotion on participants' judgements are indeed driven by the distinctiveness of emotional words compared to a background of neutral words, as predicted by Hourihan et al.'s (2017) cognitive account, participants are expected to assign higher JOLs to emotional words compared to neutral words. As described above, no differences between negative and positive valence words are expected. Similarly, it is expected that recall performance will follow this trend. Importantly, the post-test questionnaire is expected to capture participants' pre-existing beliefs regarding the differential memorability of emotional and neutral information.

2.2 Method

2.2.1 Participants. Sixty-eight undergraduate students at Memorial University of Newfoundland participated for one course credit. Two participants were not included in all data analyses due to failure to follow initial instructions, leaving a total of 66 participants.

2.2.2 Materials. The experiment ran on a computer (Windows 7 OS), using E-Prime 2.0 (Psychology Software Tools, Pittsburgh, PA). Two mixed lists of 20 words were selected from the ANEW database (Bradley & Lang, 2010; see Appendix A for word lists). Words were manually selected to be from the mid-range of arousal values (approximately 5.38). One list was composed of ten neutral and ten negative valence words, the other list was composed of ten neutral and ten positive valence words. The upper half of Table 2.1 displays the descriptive data for valence and arousal ratings, and frequency values for each word in Experiment 1. The lists differed significantly on valence [neutral vs. negative words: $t(18) = 10.88$, $p < .001$, neutral vs. positive words: $t(18) = 17.36$, $p < .001$, negative vs. positive words: $t(18) = 28.65$, $p < .001$] but not on arousal [neutral vs. negative words: $t(18) = 1.07$, $p = .299$, neutral vs. positive words: $t(18) = 0.19$, $p = .854$, negative vs. positive words: $t(18) = 0.922$, $p = .369$]. SUBTLEXus log word frequencies were used as a measure of word frequency (Brysbaert & New, 2009). The words did not differ significantly on log word frequency, neutral vs. negative words: $t(18) = 0.52$, $p = .610$, neutral vs. positive words: $t(18) = 0.98$, $p = .342$, negative vs. positive words: $t(18) = 0.28$, $p = .780$.

The post-test questionnaire was included to query participants about their explicit beliefs regarding the memorability of emotional words (see Appendix B). The first several questions asked them about their thoughts about the experiment in general. For example, the first question stated, "In the experiment, do you think you recalled some words more easily than others?"; the second question stated, "Do you think your predictions of future recall (JOLs) were accurate?". Other questions explicitly asked for their thoughts and beliefs regarding emotion and memory in general. For example, "Do you believe emotional words are remembered differently than neutral words?". As a measure of mood congruency, the Positive and Negative Affect Schedule (PANAS; Watson, Clark & Tellegen, 1988) was included as the final question on the questionnaire. This measure was chosen based on its high internal consistency and representativeness of both positive and negative affect.

2.2.3 Design. Emotion (Emotional vs. Neutral) and List Context (Negative vs. Positive) were manipulated within-subjects. Mean JOLs and mean proportion recall were the dependent measures obtained. The correspondence between JOLs and recall for each participant was computed with gamma, and the mean value for each word list was obtained. In order to assess the relative strength of support for each hypothesis, the results of the JOL and recall analyses were subjected to Bayesian analysis. The posterior probabilities were computed in Microsoft Excel by inputting the appropriate information from the SPSS ANOVA output (see Masson, 2010). Questionnaire responses were coded based on the scheme in Appendix C. Participants' scores on both the negative and positive subscales of the PANAS were computed as the sum of values assigned to each item. Further, a negative and positive emotional benefit score for each participant was computed by subtracting the mean proportion of neutral items from the mean of emotional items in the respective list (for both JOLs and recall performance). Correlational analyses were then conducted between participants' PANAS scores and their computed benefit scores.

2.2.4 Procedure. Participants were tested in a room where up to three participants could be tested simultaneously. Each participant completed two blocks consisting of a study phase, a distractor phase, and a recall test phase. Participants were instructed that they would be shown two different word lists to study; each word would be presented one at a time, and their goal would be to remember as many words as possible for a subsequent memory test. Each word list was composed of an equal number of neutral and emotional words, such that participants studied one list of positive and neutral valence words, and another list of negative and neutral valence words.[1] Each list was composed of 20 words, such that all participants saw a total of 40 words. Word list order was counterbalanced across participants (i.e., positive and neutral word list first, or negative

[1] Neutral valence words were not repeated across lists

and neutral word list first) and word presentation order was randomized without restrictions (e.g., random presentation of emotional and neutral words in each list).

Each study trial began with a 500ms blank screen. Words were individually presented in 18 pt. Bell MT font (black on a white background) for 2000ms and were followed by another blank screen for 500ms. After each word, participants were asked to make a prediction of how likely they thought it was that they would remember the word during a later recall test (i.e., make a JOL). Responses were selected from a scale presented at the bottom of the screen, which ranged from 1 ("definitely WILL NOT remember") to 7 ("definitely WILL remember"). JOLs were self-paced.

Following the study phase, participants were asked to complete a distractor task for three minutes. This task required participants to determine which of two alternatives was the same as the target geometric image, but rotated, by pressing the appropriate key on the keyboard. They were then given instructions for the test phase. Participants were asked to recall as many of the words from the most recent study phase that they could remember, in any order. They were instructed to type the words one at a time and press 'ENTER' to submit each word. Recall was self-paced, and participants were instructed to guess when uncertain. An additional distractor task of simple arithmetic was presented after the first recall test for one minute. The second block followed the same procedure as the first, but a different word list was presented.

Once participants had completed both blocks, they were instructed to see the researcher for the next part of the experiment. Participants were then asked to complete an eight item paper-and-pencil questionnaire (see Appendix B), and were given as much time as required for completion. The entire procedure lasted approximately 45 minutes.

Table 2.1 *Characteristics of the Word Lists in Experiments 1 and 2*

	Valence (1-9)		Arousal (1-9)		Log Word Frequency (number per million)	
	Range	**Mean**	**Range**	**Mean**	**Range**	**Mean**
Experiment 1						
Neutral (negative list)	[3.97, 5.30]	4.69 (.43)	[2.92, 6.38]	4.51 (.97)	[1.48, 3.73]	2.76 (.84)
Neutral (positive list)	[4.39, 5.43]	4.91 (.35)	[3.50, 5.43]	4.55 (.68)	[1.42, 3.05]	2.21 (.60)
Negative	[1.61, 3.00]	2.61 (.42)	[4.13, 5.42]	4.87 (.45)	[1.42, 3.56]	2.58 (.71)
Positive	[7.14, 8.26]	7.75 (.38)	[2.87, 5.64]	4.61 (.78)	[1.36, 3.39]	2.49 (.68)
Experiment 2						
Neutral	[4.00, 6.31]	5.14 (.55)	[2.95, 6.93]	4.84 (.97)	[1.23, 3.76]	2.62 (.56)
Negative	[1.25, 3.76]	2.66 (.61)	[3.34, 6.84]	4.99 (.76)	[1.51, 4.43]	2.59 (.61)
Positive	[6.73, 8.43]	7.55 (.42)	[2.97, 6.77]	4.84 (.84)	[1.54, 3.83]	2.75 (.59)

Note: Standard deviations are displayed in parentheses beside their respective means.

2.3 Results

Statistical significance was measured at $p < .05$ in all of the analyses to be discussed. Data from two participants were eliminated from all analyses due to failure to follow instructions (e.g., not completing the recall task), leaving 66 participants. Data from ten of these participants were not included in the gamma analysis due to missing data (e.g., if all JOLs are equal in a given condition, gamma cannot be computed); therefore, 56 participants were included in this analysis. Data from one participant was not included in correlational analyses due to failure to follow instructions (e.g., not assigning values to all items on the PANAS), leaving 65 participants for these analyses.

2.3.1 JOLs. A 2 (Emotion: Emotional vs. Neutral) x 2 (Context: Negative List vs. Positive List) x 2 (List Order: Negative First vs. Positive First) mixed factors analysis of variance (ANOVA) with emotion and list context as within-subjects factors and list order as a between-subjects factor was conducted. As there was no significant effect of list order, or any interactions with emotion or context, list order was collapsed, all *F*'s < 1.119, all *p*'s > .294.

A 2 (Emotion: Emotional vs. Neutral) x 2 (Context: Negative List vs. Positive List) repeated measures ANOVA with both emotion and context as within-subjects factors was conducted. As hypothesized, there was a significant main effect of emotion, $F(1, 65) = 107.83$, $MSE = 0.27$, $p < .001$, $\eta_p^2 = .624$, such that participants gave higher JOLs to emotional words (both negative and positive) relative to neutral words. There was no effect of context, $F(1, 65) = 0.68$, $MSE = 0.32$, $p = .412$, suggesting that JOLs did not differ across lists. The emotion x context interaction was also non-significant,

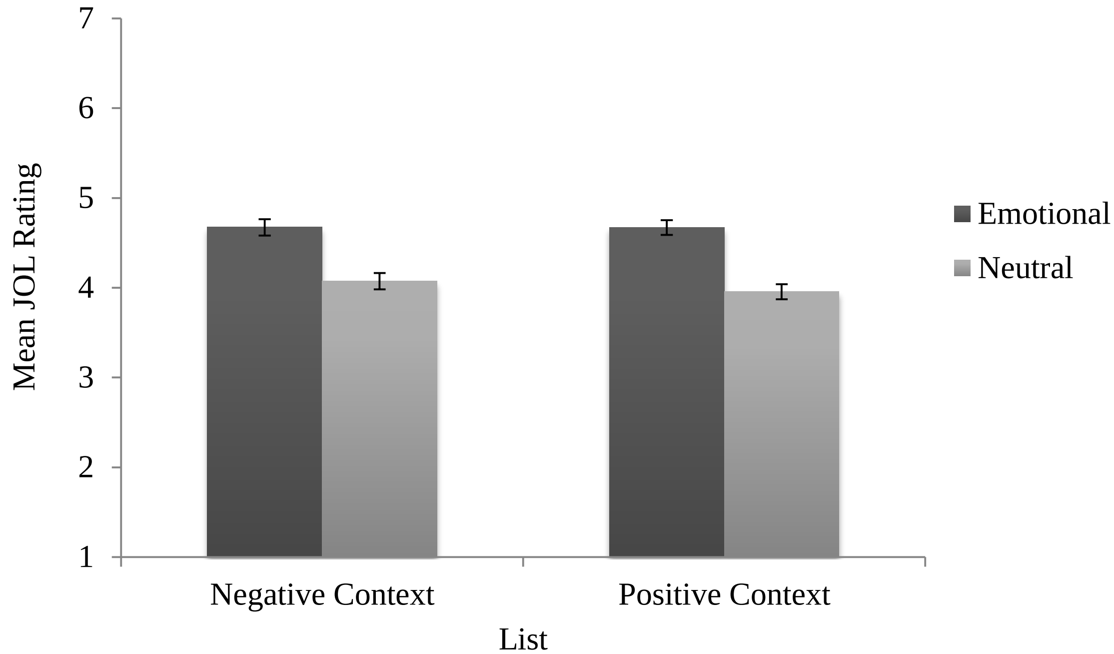

Figure 2.1 Mean JOL ratings across lists in experiment 1. Error bars represent standard error of the mean.

F (1, 65) = 1.07, *MSE*= 0.20, *p* = .305, suggesting that JOLs given to emotional words and neutral words did not differ depending on which list the words were shown in. See Figure 2.1 for mean JOLs across lists. Bayesian analyses found that the posterior probability in favor of the main effect of emotion was $p_{BIC}(H_1 \mid D) = 1.00$, providing very strong evidence that emotion influenced participants' JOLs, compared to neutral items. In terms of list context and the emotion x context interaction, there was positive evidence supporting the null, $p_{BIC}(H_0 \mid D) = .852$, $p_{BIC}(H_0 \mid D) = .826$, respectively.

2.3.2 Recall Performance. Recalled words were scored on strict criteria. Words that were spelled incorrectly but were clearly the correct item (e.g., "misqutio" instead of "mosquito") were scored correctly. Words that were studied but were recalled from the inappropriate list were scored as incorrect (e.g., words from the first list recalled during the second recall test). Words that were not studied (e.g., “scooter”), or words that were semantic variations of the studied words (e.g., “disgusting” instead of “disgusted”) were marked as intrusions. The mean number of intrusions across participants was low (M = 1.65, SEM = 0.19), and therefore was not analyzed further. See Figure 2.2 for mean recall across lists.

Similar to above, a 2 (Emotion: Emotional vs. Neutral) x 2 (Context: Negative List vs. Positive List) x 2 (List Order: Negative First vs. Positive First) mixed factors ANOVA with emotion and context as within-subjects factors and list order as a between-subjects factor was conducted. As there were no significant effects of list order, or any interactions with emotion or context (all F’s < 0.437, all p’s > .511) list order was collapsed. As such, a 2 (Emotion: Negative vs. Positive) x 2 (Context: Negative List vs. Positive List) repeated measures ANOVA with both emotion and list context as within-subjects factors was conducted. As hypothesized, there was a main effect of emotion, $F(1, 65) = 20.32$, $MSE = 0.03$, $p < .001$, $\eta_p^2 = .238$, such that emotional words (both negative and positive) were recalled more frequently than neutral words. There was no main effect of list context, $F(1, 65) = .72$, $MSE = 0.03$, $p = .399$, suggesting that recall performance did not differ based on whether the words were negative or positive in valence. The emotion x context interaction approached but did not reach significance, $F(1, 65) = 3.25$, $MSE = 0.02$, $p = .076$. To follow-up on this interaction, paired t-tests comparing recall of emotional words to the neutral words in their respective lists found that they were

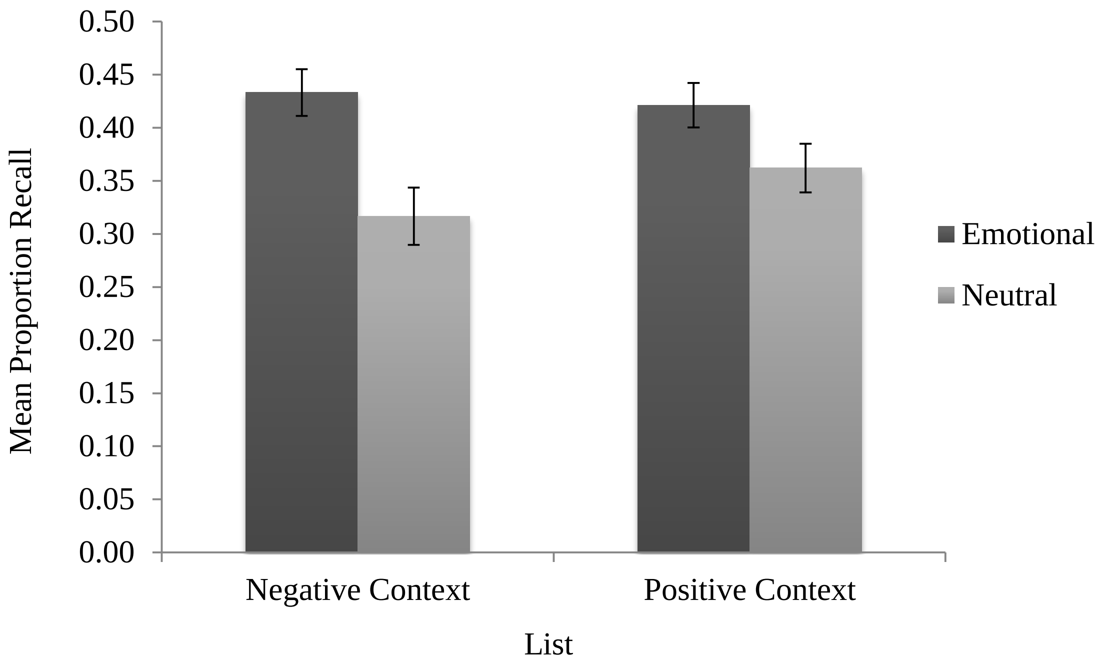

Figure 2.2 Mean proportion recall across lists in experiment 1. Error bars represent standard error of the mean.

significantly different, both p's < .01. Taken together, these results suggest that emotional words were remembered better than neutral words, and emotional words were not remembered significantly differently across lists. Additionally, Bayesian analyses showed that the resulting posterior probability in favour of the main effect of emotion was $p_{BIC}(H_1 \mid D) = .999$, providing very strong evidence that emotional words were better recalled than neutral words. In terms of the influence of list context, there was positive evidence supporting the null, $p_{BIC}(H_0 \mid D) = .850$, and weak evidence supporting the null emotion x context interaction, $p_{BIC}(H_0 \mid D) = .617$

2.3.3 Gamma. Goodman and Kruskal's gamma (G or γ) is generally used as a nonparametric measure of rank correlation. When variables are measured at the ordinal

Table 2.2 *Mean Gamma Values Across Lists in Experiment 1*

Word List	**Gamma**
Negative	.32 (.50)
Neutral (Negative)	.39 (.51)
Positive	.39 (.45)
Neutral (Positive)	.39 (.49)

Note: Standard deviations are in parentheses besides their respective means.

level, it can be used as a measure of the direction and associative strength between both variables (Nelson, 1984). Here, gamma was used as a measure of correspondence between participants' JOLs and recall performance. Gamma values were computed for each participant, and the overall values were analyzed. See Table 2.2 for mean gamma values.

Participants' predictions were found to be significantly different from zero, suggesting that they were accurate (i.e., JOLs tended to correlate with actual recall) across all lists (all *p*'s $< .001$). These results were further evaluated by using a 2 (Emotion: Emotional vs. Neutral) x 2 (List Context: Negative vs. Positive) repeated measures ANOVA. There was no effect of emotion, F (1, 55) = 0.199, MSE = 0.263, p = .657, nor was there an effect of list context, F (1, 55) = 0.205, MSE = 0.292, p = .652. The emotion x context interaction was also non-significant, F (1, 55) = 0.347, MSE = 0.202, p = .559, $p_{BIC}(H_0 \mid D) = .868$. Bayesian analyses found positive evidence supporting the null in terms of all variables: Emotion, $p_{BIC}(H_0 \mid D) = .878$, context, $p_{BIC}(H_0 \mid D) = .878$, the emotion x context interaction, $p_{BIC}(H_0 \mid D) = .868$. These results suggest that JOLs predicted free recall equivalently for emotional and neutral words across lists.

2.3.4 Questionnaire Responses. Self-report responses on the questionnaire were coded based on Appendix C. For a count of responses for each question, see Table C1. Of primary interest were questions three, five, and six; these results will be discussed in detail. Two independent raters coded the responses. There was strong agreement between the raters, Cohen's κ for all questions > .906, $p < .001$. A third independent rater resolved any disagreements.

2.3.4.1 Question 3. This question was included to ask participants if they noticed anything about the composition of the word lists. Of the 66 participants who were included in the analysis, 66.7% responded "yes", while 33.3% responded "no". Of the 66.7% who responded "yes", 68.2% spontaneously reported they noticed emotional words in the list.

2.3.4.2 Question 5. This question was included to ask participants if they noticed any emotional words in their lists. For this question, participants were simply required to respond "yes" or "no". Of the 66 participants who were included in the analysis, 92.4% indicated that they noticed emotional words, while only 7.6% indicated that they did not.

2.3.4.3 Question 6 and 6(a). For this question, participants were asked if they believe emotional words are remembered differently than neutral words. This question was the most relevant in terms of support for the hypothesis. Recall that it was proposed that the metamnemonic effect of emotional stimuli might be in part due to

Table 2.3 *Correlations Between Participants' JOLs and Recall Performance and the Positive (+) and Negative (-) Subscales of the PANAS*

	+PANAS	-PANAS	+JOL Benefit	-JOL Benefit	+Recall Benefit	-Recall Benefit
+PANAS	1.00					
-PANAS	.021	1.00				
+JOL Benefit	.213	.160	1.00			
-JOL Benefit	.229	.179	.142	1.00		
+Recall Benefit	-.304*	.149	.333**	.128	1.00	
-Recall Benefit	.236	.188	.283*	.365**	.205	1.00

Note: * denotes $p < .05$; ** denotes $p < .01$.

participants' explicit beliefs regarding the memorability of emotional words. As expected, the vast majority held this belief: 90.9% responded "yes", while only 9.1% of participants responded "no". Of those that responded "yes", 51.7% indicated that emotional words are better remembered than neutral words due to their relationship with a memory. For example, one participant explained that a negative word could be related to a significant negative event in one's life (e.g., death), whereas other [neutral] words have no significance.

2.3.4.4 The PANAS. Correlational analyses between participants' positive and negative benefit scores (i.e., mean emotional minus mean neutral JOL and recall performance) and their scores on the two PANAS subscales (i.e., sum of responses for each subscale) yielded a significant negative correlation between recall of the positive list and scores on the positive PANAS subscale, $r(54) = -.304$, $p < .05$. No other correlations were significant (for both recall performance and JOLs). See Table 2.3 for a summary of all correlations.

2.4 Discussion

In order to examine how list structure influences the metamnemonic effect of emotion, participants were asked to study mixed lists of emotional and neutral words, and then make immediate JOLs after each word. After studying each word list, they were asked to complete a self-paced free recall test. It was expected that participants would be able to explicitly notice the emotional words as being emotional due to the "background" of neutral words, and subsequently assign higher JOLs to these items. Recall performance was expected to follow this trend. Participants were also asked to complete a post-test questionnaire regarding their thoughts about the experiment, and their beliefs about emotion and memory in general. This measure was expected to show that all participants held a belief that emotional information was more memorable than non-emotional information, providing support for a cognitive account of the metamnemonic benefit of emotion.

As expected, participants gave higher JOLs to emotional words (both negative and positive) compared to neutral words, and their recall performance followed this pattern. Replicating Hourihan et al. (2017), the effect of emotion on participants' metamnemonic performance was evident even though all words were equated on the dimension of arousal. Of most interest is the finding that participants' JOLs were sensitive to differences in valence. Because this difference emerged in spite of the words being equated on arousal, this result provides evidence against a physiological account of emotional effects on metamemory. Participants' sensitivity to the influence of valence on memory performance is instead more consistent with the idea that the effects of emotional words on metamemory are influenced by intentional strategies. Hourihan et al. (2017) suggested that this effect only occurs when participants are able to explicitly notice the emotional words when compared to a 'background' of neutral words in the same list. The results from Experiment 1 are congruent with this proposition.

Results from the questionnaire also support a relationship between participants' JOLs and their beliefs. Nearly all participants reported noticing that there were emotional words in their study lists, and indicated that they held a belief that emotion influences memory. My interpretation is that participants explicitly notice emotional items when compared to neutral items in a mixed-list context, and incorporate a pre-existing belief regarding the items' differential memorability, which drives the emotional JOL effect. Additional results from the PANAS indicate that the results observed in Experiment 1 were not likely driven by participants' current mood. If mood congruency were indeed a factor that contributed to participants' performance, it would be expected that participants who scored higher on one of the subscales (either positive or negative) would show a larger recall benefit for the congruent items (either positive or negative). Surprisingly, the only significant result was a negative correlation between recall of positive items and the positive PANAS subscale. It appears as though participants who reported being in a positive mood actually recalled f*ewer* positive valence words. Mood-incongruence is driven by individual differences, such as individuals' self-esteem, world-views, and attitudes. For example, Berkowitz (2000) has shown that individuals scoring higher on levels of self-esteem recall more positive information, even following a negative mood induction. Detailed participant demographics were not collected in Experiment 1, therefore it is entirely possible that this correlation is indeed a factor of participants' attitudes; however, given that a mood induction was not involved in the procedure, nor were any other correlations between a performance benefit and subscale score found to be significant, this relationship is most likely spurious.

Moreover, it remains unknown as to whether participants will show a similar trend in a pure-list context. That is, will participants still assign higher JOLs to emotional words relative to neutral words when these items are not clearly contrasted? Experiment 2 was conducted to further examine this scenario.

3.1 Experiment 2

In Experiment 1, it was found that participants gave higher JOLs to emotional words (both positive and negative) compared to neutral words in the same list, and recall performance followed this trend. These results were not due to mood-congruency effects, and results on the questionnaire suggest that participants believe that emotional information is more memorable than non-emotional information.

Based on the idea that the metamnemonic effect of emotion depends on the distinctiveness of emotional items (relative to a background of neutral items), the cognitive account predicts that a design that eliminates this relative distinctiveness would eliminate any metamnemonic effect of emotional words. Simply put, effects should be limited to mixed-list designs. In Experiment 2, a pure-list design (lists composed of words equated for valence) was utilized to examine this possibility. Here, participants were asked to study lists that were composed of either negative, positive, or neutral words, make immediate JOLs after each word, and complete a self-paced recall test. In Experiment 1, emotional and neutral words were directly contrasted in a mixed list, which appears to have led participants to notice emotional words as distinctive, relative to neutral words. Subsequently, they assigned higher JOLs to emotional words, and reported having a pre-existing belief that emotion benefits memory. Since participants were asked to study only one type of word in Experiment 2, it was predicted that there would be no effect of emotion between participants. That is, due to the fact that participants studied pure word lists (positive, negative, or neutral valence only), it was expected that they would be less likely to notice positive or negative words as "emotional", and therefore be much less likely to consciously assign higher JOLs to these words. Based on these methodological differences, it was hypothesized that fewer participants would report noticing emotional words and use this distinctiveness strategically (relative to Experiment 1), but it was expected that they would be just as likely to hold the belief that emotion affects memory.

3.2 Method

3.2.1 Participants. Ninety undergraduate students from Memorial University of Newfoundland participated for one course credit or $10 CAD. Participants were pseudo-randomly assigned to word-list condition based on their arrival to the laboratory (i.e., Negative, Positive, or Neutral List). There were 30 participants in each condition.

3.2.2 Materials. Materials were similar to Experiment 1, save the words presented to participants. Three lists of words were selected from the ANEW database (Bradley &

Lang, 2010; see Appendix A for word lists). The bottom half of Table 2.1 displays mean valence and arousal ratings, and mean word frequency for the three lists in Experiment 2. The lists differed significantly on valence [neutral vs. negative words: $t(78) = 19.05$, $p < .001$, neutral vs. positive words: $t(78) = 22.02$, $p < .001$, negative vs. positive words: $t(78) = 41.84$, $p < .001$] but not on arousal [neutral vs. negative words: $t(78) = 0.76$, $p = .451$, neutral vs. positive words: $t(78) = 0.02$, $p = .985$, negative vs. positive words: $t(78) = 0.80$, $p = .424$]. The words also did not differ significantly in terms of log word frequency [neutral vs. negative words: $t(78) = 0.21$, $p = .838$, neutral vs. positive words: $t(78) = 1.03$, $p = .308$, negative vs. positive words: $t(78) = 1.18$, $p = .241$].

3.2.3 Design. Emotion (Negative vs. Positive vs. Neutral) was manipulated between-subjects, and List Number (First vs. Second) was manipulated within-subjects. Mean JOLs and mean proportion recall were the dependent measures obtained. The correspondence between JOLs and recall for each participant was computed with gamma, and the mean value for each word list was obtained. Additional Bayesian analyses were conducted as per Experiment 1 to permit quantification of evidence for the Null hypothesis (see Masson, 2010). In correspondence with Experiment 1, questionnaire responses were coded based on the scheme in Appendix C. PANAS scores were summed for each participant and correlational analyses were conducted with JOLs and recall performance (i.e., benefit scores, as per Experiment 1). To determine whether responses on the questionnaire differed between experiments, further analyses were conducted with a chi square test of independence.

3.2.4 Procedure. The procedure of Experiment 2 was identical to that of

Experiment 1 (i.e., study, distractor, recall test), but participants were asked to study words of only one type—neutral, negative, or positive—for both of their study lists.

3.3 Results

Statistical significance was measured at $p < .05$ in all of the analyses to be discussed. Ninety participants were included in quantitative analyses. Data from four participants were not included in correlational analyses due to failure to follow instructions (e.g., not assigning values to all of the items in the PANAS), leaving 86 participants in these analyses (29 negative, 29 neutral, 28 negative).

3.3.1 JOLs. A mixed factors ANOVA was conducted with List Number (First vs. Second) as a within-subjects factor and Emotion (Negative vs. Positive vs. Neutral) as a between-subjects factor. As hypothesized, there was no effect of emotion, $F(2, 87) = 1.82$, $MSE = 1.04$, $p = .168$, suggesting that participants' predictions did not differ based on the emotion (negative, positive, or neutral) of the words they studied. There was also no effect of list number, $F(2, 87) = 3.346$, $MSE = 0.168$, $p = .071$, nor was there a list x emotion interaction, $F(2, 87) = 1.10$, $MSE = 0.167$, $p = .337$, suggesting that JOLs for emotional and neutral words did not differ as a function of whether the words were shown in the first or second list. Given the theoretical interest in the observed null difference, a power analysis was conducted to compare JOLs between emotional and neutral conditions. The observed difference of 0.183 between the negative and neutral conditions produced a δ value of 0.918. To obtain a power of 0.80, this effect would require approximately 279 participants per group to reach statistical significance (at $p = .05$). Comparing positive and neutral conditions, the observed difference of 0.354 produced a

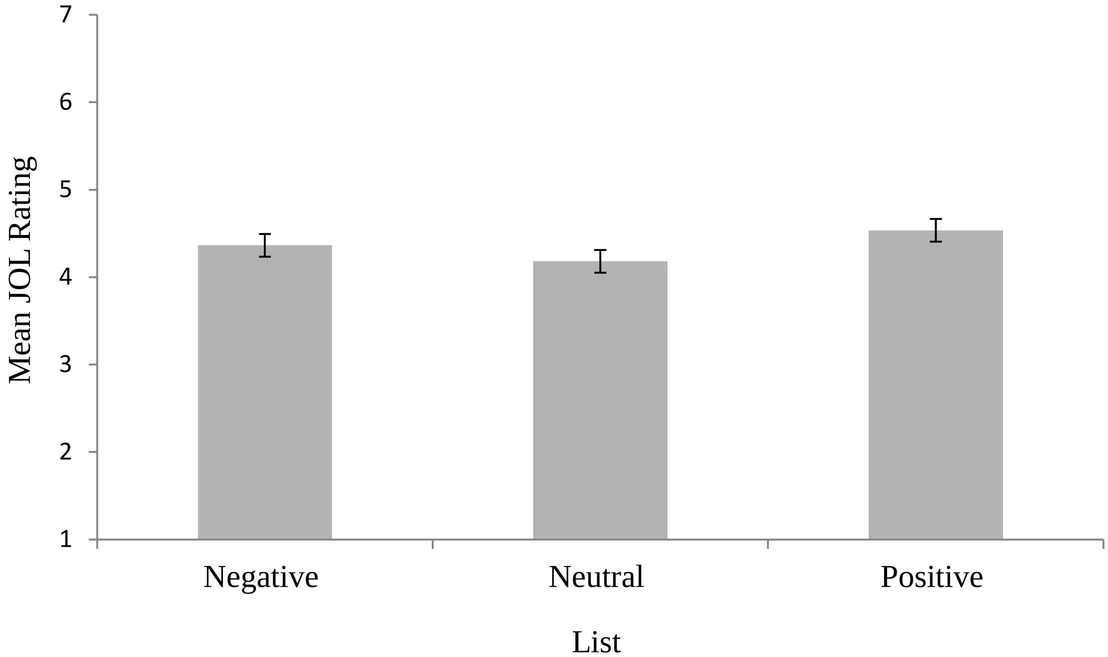

Figure 3.1 Mean JOL ratings across lists in Experiment 2. Error bars represent standard error of the mean.

δ value of 1.693. To obtain a power = 0.80, this effect would require approximately 82 participants per group to reach statistical significance (at p = .05). See Figure 3.1 for mean JOL ratings between lists.

Further, Bayesian analyses were conducted in order to assess the probability of support for the null hypothesis. Regarding the influence of emotion, the resulting posterior probability in favor of the null hypothesis was $p_{BIC}(H_0 \mid D) = .934$, showing positive evidence that JOLs did not differ based on the emotion of the words that participants studied. There was weak support favouring the null with respect to list number, $p_{BIC}(H_0 \mid D) = .635$, and strong evidence supporting the null interaction, $p_{BIC}(H_0 \mid D) = .967$.

3.3.2 Recall Performance. Recall was scored similarly to that of Experiment 1. The mean number of intrusions across participants was low (*M* = 1.60, *SEM* = 0.18) and therefore was not analyzed further. Similar to above, a mixed factors ANOVA was conducted with List Number (First vs. Second) as a within-subjects factor and Emotion (Negative vs. Positive vs. Neutral) as a between-subjects factor. As hypothesized, there was no effect of emotion, $F(2, 87) = .69$, $MSE = 0.03$, $p = .503$, suggesting that recall performance did not differ based on the emotion of the words participants studied. There was a main effect of list number, $F(1, 87) = 6.78$, $MSE = 0.01$, $p = .011$, $\eta_p^2 = .072$, such that recall performance was better in the first list (*M* = 0.41, *SD* = 0.14) compared to the second list (*M* = 0.37, *SD* = 0.14). There was also no emotion x list number interaction, $F(2, 87) = 2.33$, $MSE = 0.01$, $p = .103$. Given the theoretical interest of the null results, a power analysis was conducted in order to compare recall performance between the emotional and neutral conditions. With respect to the negative and neutral conditions, the observed difference of 0.006 produced a δ value of 0.174 (observed power < .15). To obtain a power = 0.80, this effect would require over 7000 participants to reach statistical significance (at $p = .05$). Comparing positive and neutral words, the observed difference of 0.035 produced a δ value of 0.939 (observed power < .15). To obtain a power = 0.80, this effect would require approximately 267 participants to reach significance. See Figure 3.2 for mean proportion recall between lists.

Bayesian analyses showed strong support for the null hypothesis with respect to the effect of emotion, $p_{BIC}(H_0 \mid D) = .978$. There was positive support for the alternate hypothesis with respect to list number, $p_{BIC}(H_1 \mid D) = .753$. Results favoured the null

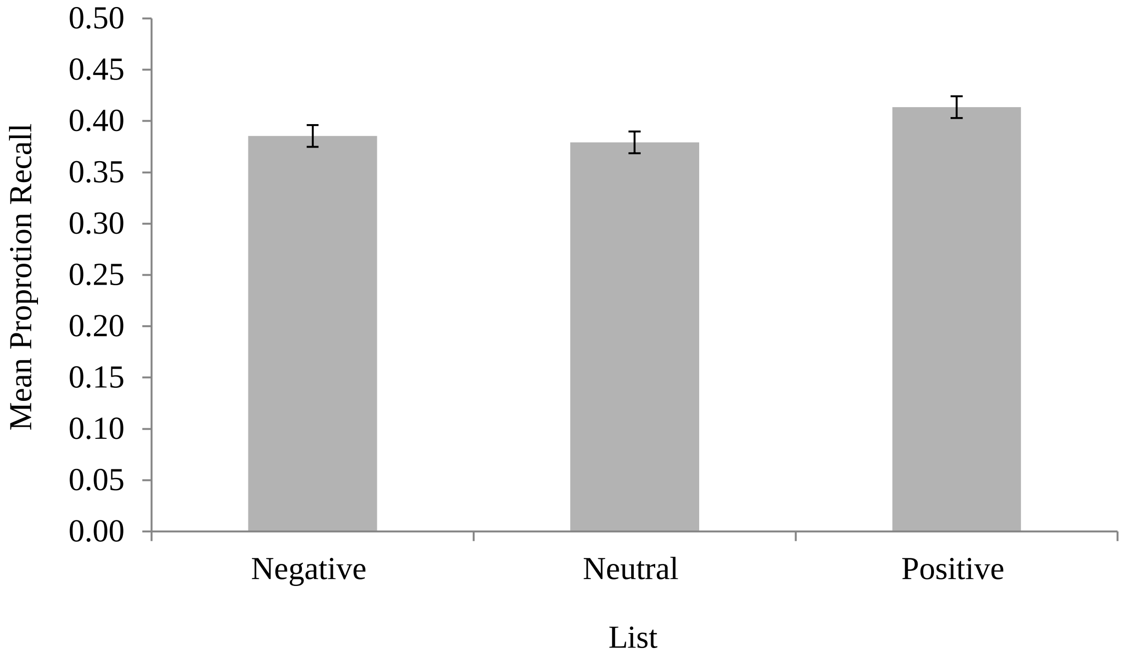

Figure 3.2 Mean proportion recall across lists in Experiment 2. Error bars represent standard error of the mean.

hypothesis with respect to the emotion x list number interaction, $p_{BIC}(H_0 \mid D) = .898$, supporting the notion that recall performance across lists did not differ based on emotion.

3.3.3 Gamma. As a measure of correspondence between JOLs and recall performance, t-tests comparing gamma values to zero showed that participants' predictions were accurate across conditions, all *p*'s < .001. To further assess participants' relative accuracy, a one-way ANOVA with gamma as the dependent measure and emotion as the independent measure was conducted. Results yielded no effect of emotion, F (2, 87) = 0.325, *MSE* = 0.054, $p = .724$. Bayesian analysis also yielded strong evidence in favour of the null, $p_{BIC}(H_0 \mid D) = .985$, suggesting there were no differences in participants' accuracy based on which emotional words (negative, positive, or neutral)

Table 3.1 *Mean Gamma Values Across Lists in Experiment 2*

Word List	**Value**
Negative	.30 (.24)
Positive	.31 (.22)
Neutral	.35 (.24)

Note: Standard deviations are in parentheses besides their respective means.

they studied. See Table 3.1 for mean gamma values.

3.3.4 Questionnaire Responses. Consistent with the first experiment, the same two raters coded responses in Experiment 2 based on the coding scheme set in Appendix C. Of primary interest were questions three, five, and six; participants' overall responses on these questions will be discussed in detail. For a count of responses for each question, see Table C2. Cohen's κ was used as a measure of inter-rater reliability for each question. There was strong agreement between the raters, κ's for all questions > .914, $p < .001$. A third independent rater solved any disagreements.

3.3.4.1 Question 3. This question was included to ask participants if they noticed anything about the composition of the word lists. Of the 90 participants who were included in the analysis, 63.3% responded "yes", while 36.7% responded "no". Of the 63.3% who responded "yes", 49.1% noted that they noticed semantic relationships between words. For example, one participant [incorrectly] noted that many of the words began with the same letter. More interestingly, of the 63.3% who responded "yes", 43.9% spontaneously reported they noticed emotional words in the list. This 43.9% represents

56.7% of those who studied the negative list, and 26.7% of those who studied the positive word list. No participants who studied the neutral list spontaneously reported noticing emotional words.

3.3.4.2 Question 5. Question five was included to ask participants if they noticed any emotional words in their lists. For this question, participants were simply required to respond “yes” or “no”. 82.2% of participants indicated that they noticed emotional words, while 17.8% indicated that they did not. Specifically, all of the participants who studied negative words reported that they noticed emotional words, 96.7% of those who studied positive words noticed emotional words, and 50% of those who studied neutral words claimed that they noticed emotional words. This finding will be explained in the discussion.

3.3.4.3 Question 6 and 6(a). Question six was included to ask participants if they believed emotional words are remembered differently than neutral words. As expected, the majority of participants held this belief: 92.2% responded “yes”, while only 7.8% of participants responded “no”. Of those that responded “yes”, 55.4% indicated that emotional words are better remembered than neutral words due to their relationship with a memory. For example, a participant noted that it was easier to recall words if they had been associated with an emotional experience. More specifically, all participants who studied negative words reported they held this belief, 90% of participants who studied the positive words held this belief, and 86.7% of participants who studied the neutral word list held this belief.

3.3.4.4 Chi square test of independence. A chi-square test of independence was performed to determine whether participants in Experiments 1 and 2

Table 3.2 *Correlations Between Participants' JOLs and Recall performance in the Negative List of Experiment 2 and the Positive and Negative Subscales of the PANAS*

	Mean Recall	Mean JOLs	Positive PANAS	Negative PANAS
Mean Recall	1.00			
Mean JOLs	.063	1.00		
Positive PANAS	.137	-.003	1.00	
Negative PANAS	.000	-.233	-.142	1.00

differed in their responses on questions 3, 5, and 6 of the questionnaire. All participants in who studied the neutral word lists in Experiment 2 were not included in any of the following analyses, as they did not study emotional items. There were no differences in participants spontaneously noticing emotional words between experiments, χ^2 (1, N = 126) = 0.183, p = .183. There were no differences in participants noticing emotional words when explicitly asked, χ^2 (1, N = 126) = 2.42, p = .120. There were also no differences in participants reporting a belief that emotional information was remembered differently than neutral information, χ^2 (1, N = 126) = 0.793, p = .373.

3.3.4.5 The PANAS. The correlation between mean JOLs and scores on the positive PANAS was marginally significant in the positive list context, p = .058; however, no other correlations were significant, all p's > .122. Taken together, these

Table 3.3 *Correlations Between Participants' JOLs and Recall performance in the Positive List of Experiment 2 and the Positive and Negative Subscales of the PANAS*

	Mean Recall	Mean JOLs	Positive PANAS	Negative PANAS
Mean Recall	1.00			
Mean JOLs	.322	1.00		
Positive PANAS	.009	.356	1.00	
Negative PANAS	.001	.278	.307	1.00

Table 3.4 *Correlations Between Participants' JOLs and Recall performance in the Neutral List of Experiment 2 and the Positive and Negative Subscales of the PANAS*

	Mean Recall	Mean JOLs	Positive PANAS	Negative PANAS
Mean Recall	1.00			
Mean JOLs	-.114	1.00		
Positive PANAS	-.027	.293	1.00	
Negative PANAS	-.155	-.214	.185	1.00

results suggest that participants' performance was not due to mood-congruency effects. See Tables 3.2, 3.3, and 3.4 for a summary of correlations in each list context.

3.4 Discussion

In Experiment 2, participants were asked to study a pure list of words (negative, positive, or neutral), make immediate JOLs after each word, and complete a free recall test after studying each list. After this, they completed a questionnaire identical to that of Experiment 1. As per the cognitive account (Hourihan et al., 2017), without the relative distinctiveness between emotional and neutral items, it was hypothesized that participants would not be able to utilize an explicit belief regarding the superior memorability of emotional words; therefore, it was expected that participants' JOLs and recall performance would not differ between lists.

As expected, there were no significant differences with respect to participants' JOLs between lists. That is, participants who studied an emotional list (negative or positive) did not assign JOLs differently than participants who studied the neutral list. Similarly, recall performance did not differ between participants. However, results yielded an unexpected significant effect of list number on recall. Better recall performance in List 1 compared to List 2 may simply be due to fatigue or proactive interference (see Keppel & Underwood, 1962; Underwood, 1957). For example, if participants judged the items to be semantically related to each other, or if participants were unable to separate the lists into separate contexts, items from the first list could have interfered with items from the second list, leading to the observed poorer recall for List 2 compared to List 1.

Importantly, participants' responses on the questionnaire support the proposition set forth in the cognitive account (Hourihan et al., 2017): Participants have a belief that emotional information should be more memorable than neutral information. Data analyses suggest that this belief was evident across both experiments; it appears as though individuals may make JOLs based on an explicit belief regarding how the intrinsic qualities of experimental stimuli are likely to affect memory. This belief, coupled with a design that obviates the distinctiveness of emotional items, appears to be the driving mechanism responsible for the metamnemonic effect of emotional words.

4.1 General Discussion

A primary goal of this thesis was to examine the untested components of the cognitive account of the emotional JOL effect proposed by Hourihan et al. (2017). In a series of experiments, they found that participants were more likely to give higher JOLs to negative valence words relative to neutral words in the same list. As such, no differences were found utilizing a pure list where words varied in both arousal and valence. Hourihan et al. suggested that the metamnemonic effect of emotion is likely limited to mixed-list designs, and is also coupled with participants' explicit beliefs regarding the memorability of emotional information, such that participants assign higher JOLs to emotional words (relative to neutral words) because they believe emotional words are more memorable. The current thesis further examined the metamnemonic effect of emotion (without the influence of arousal), and found support for the cognitive account by explicitly querying participants regarding their beliefs about the memorability of emotional information.

Two experiments were conducted wherein the emotional dimension of valence was manipulated while arousal was held constant. Participants were required to study lists of words, make immediate judgements of learning (JOLs) after each word, and complete a free recall test after each word list. In Experiment 1, participants were presented with mixed lists of emotional and neutral words, whereas participants only studied words of one valence (either negative, positive, or neutral) in Experiment 2. In Experiment 1, participants recalled more emotional words (both negative and positive) relative to neutral words. Importantly, the mixed-list design of discrete emotional and neutral categories appears to have allowed the emotional words to become relatively distinctive, as participants also indicated that the emotional words would be better recalled than the neutral words. Additionally, participants were found to be accurate in their predictions; on average, participants recalled the items they judged they would recall, and did not recall the items they judged they would not recall (regardless of emotion or list context). In Experiment 2, a pure-list design appears to have eliminated the distinctiveness between emotional and non-emotional words, eradicating the metamnemonic (and mnemonic) effect of emotion that was found in the first experiment.

Given that the trend in superior memorability for emotional information relative to neutral information is commonly replicated (see Kensinger, 2009), the finding that participants recalled more negative and positive words relative to neutral words in Experiment 1 is relatively unsurprising. What is of more interest, however, is the finding that participants assigned higher JOLs to emotional words, but only in the mixed-list design of Experiment 1. Such findings have been noted in the literature with discrete-level versions of the mixed-list design, in which a study list contains only two or three ‘types’ of items; emotional words are able stand out from neutral words as they are categorically distinctive (e.g., Hourihan et al., 2017; Talmi, 2013). When the distinctiveness of stimuli is reduced at study by presenting pure lists rather than mixed lists, the immediate mnemonic benefit for emotional words is eradicated (e.g., Dewhurst & Parry, 2000; Hourihan et al., 2017; Talmi & Moscovitch, 2004). This evidence suggests that if the emotional and neutral words were not evident as different ‘types’ of words at encoding, JOLs for emotional and non-emotional words would no longer differ because there would no longer be differences in distinctiveness.

Distinctive items are simply those items that stand out from other items (see Schmidt, 1991). Primary distinctiveness refers to items that are distinctive due to context, whereas secondary distinctiveness refers to items that stand out relative to other items due to their inherent qualities. The two experiments outlined above give precedence to the notion of primary distinctiveness. In Experiment 1, words were studied in mixed lists such that emotional items could stand out from a background of neutral items. In Experiment 2, emotional words were studied in the absence of such background items. As mentioned above, the effects of emotional content on JOLs were only found in the first experiment; comparing distinctive items to other ‘types’ of items allowed these items to be preferentially remembered (see Talmi, 2013). As such, the results of the current thesis further bolster the idea that JOLs for emotional words are a product of list composition (i.e., primary distinctiveness) rather than inherent emotional content (i.e., secondary distinctiveness).

This cognitive strategy lends itself from other memory phenomena, such as the Generation Effect (Slamecka & Graf, 1978) and the Orthographic Distinctiveness effect (e.g., Hunt & Elliot, 1980). Research generally converges on the notion that items that are in some manner distinctive from other items incur a mnemonic benefit. For example, the production effect is based on the distinctiveness of produced items relative to unproduced items (MacLeod et al., 2010; Ozubko & MacLeod, 2010; Ozubko, Major, & MacLeod, 2014). This relative distinctiveness has been attributed to design differences, such that the effect is limited to within-subjects designs (but see Fawcett, 2013). Interestingly, McDaniel and Bugg (2008) have proposed an item-order framework to account for differences in free recall performance in mixed-list designs versus pure-list designs. They suggest that the mnemonic advantage of items in mixed lists compared to pure lists is due to differential encoding of item information due to the 'type' of stimulus. More specifically, their account assumes that prior knowledge regarding item order mediates performance on free recall tasks, that list composition mediates the influence of item order of relatively distinct items, and 'unusual' items are given more item-level elaboration, regardless of list context. These propositions mirror those presented in the current thesis; however, in the words of McDaniel and Bugg, 'unusual' items are preferentially encoded, *regardless of list context.* The current experiments have shown that the metamnemonic effect of emotion is moderated by the list context, and that performance on these tasks is not affected by list order. Clearly, the distinctiveness of emotional items relative to neutral items is determined by list composition, rather than due to solely the inherent qualities of different 'types' of items.

Evidence increasingly indicates that valence can influence the memorability of emotional events even without the influence of arousal (e.g., Hourihan et al., 2017). For example, Mickley and Kensinger (2008) found that the manner in which negative and positive information is remembered is due to differences in the processes recruited during encoding (compared to the encoding of neutral information). However, this does not negate the fact that arousal influences memory, specifically, via factors that act during encoding and those that modulate memory consolidation (Kensinger, 2009). Indeed, evidence from neuroimaging studies supports the notion that emotional stimuli engage specific cognitive and neural mechanisms that enhance explicit memory beyond that of non-emotional stimuli. For example, increased brain activity in the amygdala during encoding has been linked with superior performance on delayed free recall tasks for negative information relative to neutral information (Cahill, 1996; 2004). Hamann (2001) examined the relationship between brain activity at encoding and later memory for emotional and non-emotional picture stimuli. For emotional stimuli, bilateral amygdala activity during encoding was related to better recognition memory for emotional stimuli compared to neutral stimuli; however this trend was not observed during recall. Unsurprisingly, amygdala activity was found to modulate activity in the hippocampus. From this, Hamann concluded that the amygdala plays a role in the enhancement of memory for emotional stimuli via its moderating effects on hippocampal activation. Provided, it is evident that arousal clearly does influence memory, but it does not necessarily contribute to the *metamnemonic* effect. Arousal's effects on metamemory are instead likely to be the result of its interaction with valence (e.g., Tauber & Dunlosky, 2010; Zimmerman & Kelley, 2012), and the fact that high arousal words are distinctive relative to low arousal words in the same study context. That is, participants may be operationalizing the label of "emotional" based on an aggregation of both valence and arousal. From this, it can be postulated that the dimension of arousal is sufficient, but not necessary for the observed metamnemonic effects of emotion. Additionally, emotional items that are aggregates of valence and arousal may be subject to participants' explicit beliefs regarding the memorability of emotional stimuli; as previously stated, evidence from the current thesis suggests that the metamnemonic effect relies upon the relative distinctiveness of emotional items compared to "less emotional" items, regardless of the dimension of arousal.

Moreover, if participants made JOLs based only on an interpretation of the subjective physiological response associated with reading an emotional word, no differences in JOLs should have been observed in the current experiments, because all items were equated on arousal. Instead, it seems likely that the higher JOLs observed for emotional stimuli arises because of the manner in which lists are composed. The mixed lists used in Experiment 1 contrasted emotional items with neutral items, increasing the saliency of the emotional content of the words and potentially allowing participants to base their judgements on an explicit belief about the memorability of emotional words. For example, considering Experiment 1, studying words with differences in valence may have led participants to consciously focus on the fact that the study list consisted of different 'types' of words; the requirement to provide JOLs for each word may have then led participants to consciously rely on this distinction when asked to provide JOLs (cf. Mueller, Dunlosky, Tauber, & Rhodes, 2014). Thus, as per the cognitive account proposed by Hourihan et al. (2017), performance based on the relative fluency of emotional items may potentially be due to a primarily cognitive (or belief-based) strategy.

In order to support this proposition, participants were explicitly queried about whether they noticed emotional words in their lists. Responses on the post-test questionnaire support a cognitive account of the metamnemonic effect of emotion. Specifically, 92.4% participants in Experiment 1 noted that they noticed emotional words, as did all participants in Experiment 2 who studied the negative word list, 96.7% of those who studied the positive word list, and 50% of those who studied the neutral word list. It was not expected that 50% of participants who studied the neutral list would report noticing emotional words (as all words were equated at the midpoint of valence, and thus emotionally "neutral"; e.g., table); however, it is likely that these responses are due to demand characteristics. Earlier in the questionnaire (question 3), participants were asked if they noticed anything about the composition of the word lists. Participants who studied the emotional lists spontaneously reported noticing emotional items; however, *none* of the participants who studied the neutral lists commented on the emotionality of words; instead they discussed the semantic relationships they made between words. Interestingly, when they were later explicitly asked about the emotional content of the word lists (question 5), 50% of participants who studied the neutral word lists indicated that they noticed emotional words. This substantial increase in response from question 3 to question 5 indicates that participants may have simply responded based on demand characteristics. Participants were aware that they were participating in a study regarding memory for emotional content, but were unaware of the between-subjects manipulation of emotion; therefore, their response when explicitly queried about emotional content may be due to feeling as though they "should have" noticed emotional words. Additionally, participants were asked if they noticed *any* emotional words in the lists; it is unknown as to whether participants noticed one or multiple emotional items. Importantly, data analyses yielded null effects with respect to emotion in Experiment 2, bolstering the notion that overall, these responses are likely non-reflective of participants' actual performance.

With respect to participants' beliefs regarding the differential memorability of emotional and non-emotional items, 90.1% of participants in Experiment 1 indicated that they held a belief that emotional information is more memorable than neutral information, all participants who studied the negative list in Experiment 2 held this belief, and approximately 90% of those who studied the positive and neutral lists reported this belief. Data analyses yielded no significant differences between experiments, therefore it can be postulated that responses were not due to the differences in experimental context. Coupled with the effects only found in Experiment 1, these beliefs may be the driving mechanism behind the observed metamnemonic effects of emotion.

The proposition that mnemonic effects may rely on beliefs is surprisingly not novel. Theory-based judgements (although sometimes not accurate) regarding memory have frequently been demonstrated. For example, participants' JOLs have been shown to reflect the belief that generated items are more memorable than read items (Begg et al., 1991; Slamecka & Graf, 1978), that text with illustrations is more memorable than text without (Serra & Dunlosky, 2010), and that larger font size is more memorable than smaller font size (Rhodes & Castel, 2008). Given that emotion is an intrinsic component of word stimuli, the notion that participants may also utilize pre-existing beliefs in making *predictions* about their memory for emotional stimuli is warranted (see Koriat, 1997). The interactive nature between memory monitoring and performance therefore posits that participants use list distinctiveness to categorize emotional and non-emotional words, and may then use a pre-existing belief about the differential memorability of these items to guide their judgements of future recall performance.

Dubiously, the ‘belief-driven’ metamnemonic may also be one that is subject to bias. Confirmation biases are effects in information processing that are broadly defined as the tendency to “preserve one’s existing beliefs when searching for evidence, interpreting it, or recalling it from memory” (Risen & Gilovich, 2007, p. 113). Regardless of whether participants experience stimuli while in an otherwise “neutral” state (as this study assumes), they may still selectively remember the stimuli that conform to their expectations. More specifically, although all items were chosen from a normative database (ANEW, Bradley & Lang, 2010), participants may still have interpreted a neutral word as emotional (e.g., one participant noted that the neutral word “weight” was emotional, as it could be related to issues with body weight). This theory of selective memory is relatively ambiguous, such that the literature differs with respect to mnemonic effects. For instance, schema theory predicts that information matching prior expectations will be more easily stored and recalled than information that does not match individuals’ expectations (e.g., Bartlett, 1932; Mandler, 1984; Rumelhart, 1980). If participants have explicit beliefs about the superior memorability of emotional words, this pre-existing belief may potentially bias the manner in which they encode, process, and later retrieve said information. With respect to emotion, these beliefs may be partially mediated by self-referential processing. We know that things we relate to ourselves are better recalled than things we do not relate with (e.g., Klein, Loftus, & Burton, 1989; Rogers, Kuiper, & Kirker, 1977). As stated above, one participant who studied the pure-neutral list indicated that ‘weight’ was an emotional word as it could be related to one’s body image. It is certainly possible then, that participants relate the emotional items to themselves during study; it may be this self-referential bias that allows participants to later recall the emotional items more accurately, rather than a conscious interpretation during study.

This counter-explanation, although having surface validity, does not seem likely. Given the aforementioned assumptions, it would then be expected that the majority of participants would process emotional words in this manner, which would lead us to hypothesize that emotional effects on metamemory would occur regardless of list structure. That is, if participants were instead recalling emotional words more than neutral words because they simply were biased towards self-referential processing, an effect of emotion should have been evident in Experiment 2 as well as Experiment 1. Considering no such differences were found, I argue that the metamnemonic effect of emotional words is likely driven by list structure, coupled with participants' pre-existing beliefs regarding the memorability of studied items.

4.1.1 Conclusion

Emotional information is generally remembered better than neutral information. Although participants tend to provide higher JOLs for emotional information compared to neutral information (Hourihan & Bursey, 2016; Nomi, Rhodes, & Cleary, 2013; Tauber & Dunlosky, 2012; Zimmerman & Kelley, 2010), these JOLs appear to be a product of list composition, rather than the product of a physiological response associated with reading an emotional stimulus (e.g., Hamann, 2001). It is known that participants' metacognitive judgements are often influenced by pre-existing beliefs about memory, rather than their actual memory capability; in two experiments, I have provided evidence that JOLs for emotional words may be primarily driven by belief. Simply put, participants have explicit beliefs about how emotion affects memory, and they use emotion as a cue when making predictions about their memory, but only in a context where it is relatively distinctive. As put eloquently by philosopher Martha Nussbaum (2001, p. 3), “[emotions] are parts, highly complex and messy parts, of reasoning itself”.

References

Arbuckle, T. Y., & Cuddy, L. L. (1969). Discrimination of item strength at time of presentation. *Journal of Experimental Psychology, 81*, 126-131. doi: 10.1037/h0027455.

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

Benjamin, A. S. (2003). Predicting and postdicting the effects of word frequency on memory. *Memory and Cognition, 31*, 297-305. doi:10.3758/BF03194388.

Berkowitz, L. (2000). *Causes and consequences of feelings*. West Nyack, N.Y.: Cambridge University Press.

Bless, H., Clore, G. L., Schwarz, N., Golisano, V., Rabe, C., & Wolk, M. (1996). Mood and the use of scripts: Does a happy mood really lead to mindlessness? *Journal of Personality and Social Psychology, 71*, 665-679. doi: 10.1037/0022-3514.71.4.665

Bradley, M. M., Greenwald, M. K., Petry, M. C. & Lang, P. J. (1992). Remembering pictures: Pleasure and arousal in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 379-390. doi: 10.1037/0278-7393.18.2.379.

Bradley, M. M., & Lang, P. J. (2010). Affective norms for English words (ANEW): Stimuli, instruction manual and affective ratings (Technical Report C-2). Gainesville, FL: University of Florida, Center for Research in Psychophysiology.

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

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41,* 977-990. doi: 10.3758/BRM.41.4.977.

Cahill, L., Haier, R. J., Fallon, J., Alkire, M. T., Tang, C., Wu, J., & McGaugh, J. L. (1996). Amygdala activity at encoding correlated with long-term, free recall of emotional information. *Proceedings of the National Academy of Sciences of the United States of America, 93,* 8016–8021.

Cahill, L., Haier, R. J., White, N. S., Fallon, J., Kilpatrick, L., Lawrence, C., Potkin, S. G., & Alkire, M. T. (2001). Sex-related difference in amygdala activity during emotionally influenced memory storage. *Neurobiology of Learning and Memory, 75,* 1–9. doi: 10.1006/nlme.2000.3999.

Carstensen, L. L. (1992). Social and emotional patterns in adulthood: Support for socioemotional selectivity theory. *Psychology and Ageing, 7,* 331-338.

Dewhurst, S. A., & Parry, L. A. (2000). Emotionality, distinctiveness, and recollective experience. *European Journal of Cognitive Psychology, 12,* 541-551. doi: 10.1080/095414400750050222.

Egidi, G., & Gerrig, R. J. (2009). How valence affects language processing: Negativity bias and mood congruence in narrative comprehension. *Memory & Cognition, 37,* 547-555. doi: 10.3758/MC.37.5.547.

Eich, J. M. (1980) The cue-dependent nature of state-dependent retrieval. *Memory & Cognition, 8,* 157-173. doi: 10.3758/BF03213419.

Efklides, A., & Petkaki, C. (2005). Effects of mood on students' metacognitive experiences. *Learning and Instruction, 15,* 415-431. doi: 10.1016/j.learninstruc.2005.07.010.

Efklides, A. (2016). Metamemory and affect. In J. Dunlosky & S. K. Tauber (Eds.) *The Oxford Handbook of Metamemory* (pp. 245-267). New York, NY: Oxford University Press.

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behaviour Research Methods, Instruments, and Computers, 28,* 1-11.

Fawcett, J. M. (2013). The production effect benefits performance in between-subjects designs: A meta-analysis. *Acta Psychologica*, *142,* 1-5. doi: 10.1016/j.actpsy.2012.10.001.

Fieldler, K., & Hutter, M. (2014). Memory and emotion. In T. J. Perfect & S. Lindsay (Eds.) *The SAGE Handbook of Applied Memory* (pp. 145-161). California, USA: SAGE Publications Ltd.

Ferguson, E., Moghaddam, N. G., & Bibby, P. A. (2007). Memory bias in health anxiety is related to the emotional valence of health-related words. *Journal of Psychosomatic Research*, *62,* 263-274. doi: 10.1016/j.jpsychores.2007.01.015.

Forgas, J. P. (1995). Mood and judgment: The affect infusion model (AIM). *Psychonomic Bulletin and Review, 117,* 39-66.

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

Hourihan, K. L., & Bursey, E. (2016). A misleading feeling of happiness: Metamemory for positive emotional and neutral pictures. *Memory, 25,* 35-43. doi: 10.1080/09658211.2015.1122809.

Hourihan, K. L., Fraundorf, S. H. & Benjamin, A. S. (2017). The influences of valence and arousal on judgments of learning and on recall. *Memory and Cognition, 45,* 121-136. doi: 10.3758/s13421-016-0646-3.

Hunt, R. R., & Elliot, M. J. (1980) The role of nonsemantic information in memory: Orthographic distinctiveness effects on retention. *Journal of Experimental Psychology: General, 109,* 49-74. doi: 10.1037/0096-3445.109.1.49.

James, W. (1884). What is an Emotion? *Mind, 9,* 188-205.

Kensinger, E. A. (2009). Remembering the details: Effects of emotion. *Emotion Review, 1,* 99-113. doi: 10.1177/1754073908100432.

Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory and Cognition, 31,* 1169-1180. doi: 10.3758/BF03195800.

Kensinger, E. A., O'Brien, J., Swanberg, K., Garoff-Eaton, R. J., & Schacter, D. L. (2007). The effects of emotional context on reality-monitoring performance in young and older adults. *Psychology and Aging, 22,* 752-764. doi: 10.1037/0882-7974.22.4.752.

Keppel, G., & Underwood, B. J. (1962). Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning and Verbal Behavior, 1,* 153-161. doi: 10.1016/S0022-5371(62)80023-1.

Klein, S. B., Loftus, J., & Burton, H. A. (1989). Two self-reference effects: The importance of distinguishing between self-descriptiveness judgments and autobiographical retrieval in self-referent encoding. *Journal of Personality and Social Psychology, 56*, 853-865. doi: 10.1037/0022-3514.56.6.853.

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General, 126*, 349-370. doi: 10.1037/0096-3445.126.4.349.

Koriat, A., & Bjork, R. A. (2004). Predicting one's own forgetting: The role of experience-based and theory-based processes. *Journal of Experimental Psychology: General, 153*, 643-656. doi: 10.1037/0096-3445.133.4.643.

Koriat, A., & Nussinson, R. (2009). Attributing study effort to data-driven and goal-driven effects: Implications for metacognitive judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 1338-1343. doi: 10.1037/a0016374.

MacLeod, C., & Matthews, A. (2004). Selective memory effects in anxiety disorders: An overview of research findings and their implications. In D. Resiberg & P. Hertel (Eds.), *Memory and Emotion (*pp. 155-185). New York: Oxford University Press.

MacLeod, C. M., Gopie, N., Hourihan, K. L., Neary, K. R., & Ozubko, J. D. (2010). The production effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 671-685. doi: 10.1037/a0018785.

Mandler, J. M. (1984). *Stories, scripts, and scenes: Aspects of schema theory.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Masson, M. E. (2011). A tutorial on a practical bayesian alternative to null-hypothesis significance testing. *Behavior Research Methods, 43,* 679-690. doi: 10.3758/s13428-010-0049-5.

Mather, M. (2007). Emotional arousal and memory binding: An object-based framework. *Perspectives on Psychological Science*, *2*, 33-52. doi: 10.1111/j.1745-6916.2007.00028.x.

McDaniel, M. A., & Bugg, J. M. (2008). Instability in memory phenomena: A common puzzle and a unifying explanation. *Psychonomic Bulletin and Review*, *15,* 237-255. doi: 10.3758/PBR.15.2.237.

McGeoch, J. A. (1942). *The psychology of human learning: An introduction.* New York: Longmans Green.

Mickley, K. R., Kensinger, E. A. (2008). Emotional valence influences the neural correlates associated with remembering and knowing. *Cognitive, Affective, and Behavioural Neuroscience, 8*, 143-152. doi: 10.3758/CABN.8.2.143.

Mickley-Steinmetz, K. R., & Kensinger, E. A. (2009). The effects of valence and arousal on the neural activity leading to subsequent memory. *Psychophysiology, 46,* 1190-1199. doi: 10.1111/j.1469-8986.2009.00868.x.

Mueller, M. L., Dunlosky, J., Tauber, S. K., & Rhodes, M. G. (2014). The font-size effect on judgments of learning: Does it exemplify fluency effects of reflect people's beliefs about memory? *Journal of Memory and Language, 70,* 1-12. doi: 10.1016/j.jml.2013.09.007.

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin, 95,* 109-133. doi: 10.1037/0033-2909.95.1.109.

Nomi, J. S., Rhodes, M. G., & Cleary, A. M. (2013). Emotional facial expressions differentially influence predictions and performance for face recognition. *Cognition and Emotion, 27,* 141-149. doi: 10.1080/02699931.2012.679917.

Nussbaum, M. C. (2001). *Upheavals of thought: The intelligence of emotions.* Cambridge, England: Cambridge University Press.

Ozubko, J. D., & MacLeod, C. M. (2010). The production effect in memory: Evidence that distinctiveness underlies the benefit. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 1543-1547. doi: 10.1037/a0020604.

Ozubko, J. D., Major, J., & MacLeod, C. M. (2014). Remembered study mode: Support for the distinctiveness account for the production effect. *Memory, 22,* 509-524. doi: 10.1080/09658211.2013.800554.

Psychology Software Tools, Inc. [E-Prime 2.0]. (2012). Retrieved from http://www.pstnet.com.

Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of perceptual fluency on affective judgments. *Psychological Science, 9,* 45-48. doi: 10.1111/1467-9280.00008.

Reisberg, D., & Heuer, F. (2004). Memory for emotional events. In D. Reisberg & P. Hertel (Eds.), *Memory and Emotion* (pp. 1-41). New York, NY: Oxford University Press.

Rhodes, M. G., & Castel, A. D. (2008). Memory predictions are influenced by perceptual information: evidence for metacognitive illusions. *Journal of Experimental Psychology: General*, *137,* 615–625. doi: 10.1037/a0013684.

Risen, J. & Gilovich, T. (2007). Informal Logical Fallacies. In R. J. Sternberg, H. L. Roediger III, & D. F. Halpern (Eds.), *Critical Thinking in Psychology* (pp. 110-130). Cambridge, England: Cambridge University Press.

Rogers, T. B., Kuiper, N. A., & Kirker, W. S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology*, *35*, 677-688. doi: 10.1037/0022-3514.35.9.677.

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro et al. (Eds.), *Theoretical Issues in Reading Comprehension*. Hillsdale, NJ: Lawrence Erlbaum.

Schmidt, S. R. (1991). Can we have a distinctive theory of memory? *Memory & Cognition*, *19,* 523–542. doi: 10.3758/BF03197149.

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 592– 604. doi: 10.1037/0278-7393.4.6.592.

Storbeck, J., & Clore, G. J. (2005). With sadness comes accuracy; with happiness, false memory: Mood and the false memory effect. *Psychological Science, 16*, 785-791. doi: 10.1111/j.1467-9280.2005.01615.x.

Susser, J. A., Mulligan, N. W., & Besken, M. (2013). The effects of list composition and perceptual fluency on judgments of learning (JOLs). *Memory and Cognition, 41,* 1000-1011. doi: 10.3758/s13421-013-0323-8.

Talmi, D. (2013). Enhanced emotional memory: Cognitive and neural mechanisms. *Current Directions in Psychological Science, 22,* 430-436. doi: 10.1177/0963721413498893.

Talmi, D., & McGarry, L. M. (2012). Accounting for immediate emotional memory enhancement. *Journal of Memory and Language, 66,* 93-108. doi: 10.1016/j.jml.2011.07.009.

Talmi, D., & Moscovitch, M. (2004). Can semantic relatedness explain the enhancement of memory for emotional words? *Memory, 32,* 742-751. doi: 10.3758/BF03195864.

Tauber, S. K., & Dunlosky, J. (2012). Can older adults accurately judge their learning of emotional information? *Psychology and Ageing, 27,* 924-933. doi: 10.1037/a0028447.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80,* 352-373. doi: 10.1037/h0020071.

Underwood, B. J. (1957). Interference and forgetting. *Psychological Review, 64,* 49-60.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measure of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54,* 1063-1070. doi: 10.1037/0022-3514.54.6.1063.

Zimmerman, C. A. & Kelley, C. M. (2010). "I'll remember this!" Effects of emotionality on memory predictions versus memory performance. *Journal of Memory and Language, 62,* 240-253. doi: 10.1016/j.jml.2009.11.004.

Appendix A: Word Lists

Words used in Experiment 1

	Number	Word	Valence	Arousal	Frequency
Neutral Valence	1992	policy	4.50	4.30	3.14
	2091	revert	4.53	4.73	1.56
	426	table	5.22	2.92	3.73
	408	square	4.74	3.18	3.21
	1112	assume	4.69	4.97	3.23
	227	industry	5.30	4.47	2.78
	2123	rush	3.97	6.38	3.20
	658	bland	4.10	4.29	1.75
	2331	tendon	4.90	4.90	1.48
	1056	admit	4.93	4.97	3.48
	307	patent	5.29	3.50	2.1
	941	plain	4.39	3.52	3.05
	2410	usage	4.93	4.90	1.43
	2022	propeller	5.43	4.43	1.90
	346	rattle	5.03	4.36	2.24
	2265	steam	5.00	4.35	2.84
	2249	spine	5.12	4.48	2.47
	1972	pin	4.50	5.17	2.92
	1326	compel	4.97	5.43	1.72

	1195	blurt	4.48	5.36	1.42
Negative Valence	368	sad	1.61	4.13	3.51
	365	rotten	2.26	4.53	2.95
	124	disgusted	2.45	5.42	1.96
	76	coffin	2.56	5.03	2.66
	411	stink	3.00	4.26	2.83
	272	measles	2.74	5.06	2.03
	885	mosquito	2.80	4.78	1.97
	169	foul	2.81	4.93	2.87
	704	crime	2.89	5.41	3.56
	1087	annoyance	2.97	5.18	1.42
Positive Valence	1951	peaceful	7.77	2.87	2.76
	468	vacation	8.16	5.64	3.22
	246	kindness	7.82	4.30	2.66
	35	beauty	7.82	4.95	3.39
	347	refreshment	7.44	4.45	1.72
	466	useful	7.14	4.26	2.83
	135	easygoing	7.20	4.30	1.36
	105	delight	8.26	5.44	2.21
	2239	soothe	7.98	4.96	1.83
	372	satisfied	7.94	4.94	2.90

Words used in Experiment 2

	Number	Word	Valence	Arousal	Frequency
Neutral Valence	2050	react	5.63	5.87	2.56
	928	passage	5.28	4.36	2.60
	1802	loud	4.77	6.10	2.61
	825	item	5.26	3.24	2.80
	57	butter	5.33	3.17	3.02
	1888	newspaper	5.52	4.42	3.08
	2219	smudge	4.00	4.43	1.72
	2365	ton	4.68	4.33	2.58
	2307	swipe	4.86	4.33	1.92
	2102	road	5.53	4.9	3.76
	2297	supply	5.27	4.38	2.98
	2349	thud	4.47	5.20	2.00
	2004	position	5.74	5.10	3.57
	1779	ledge	4.13	5.35	2.26
	2156	seek	5.59	5.12	2.97
	1096	application	4.69	5.62	2.71
	874	metal	4.95	3.79	3.00
	1232	bustle	5.43	5.80	1.63
	1977	plate	5.30	4.00	3.12
	410	startled	4.50	6.93	2.10

	1689	hire	5.77	5.63	3.11
	380	seat	4.95	2.95	3.60
	309	pencil	5.22	3.14	2.70
	810	indifferent	4.61	3.18	1.77
	325	prairie	5.75	3.41	2.16
	2250	sponge	5.11	3.52	2.53
	1054	adjust	5.25	5.39	2.46
	1261	cave	4.78	5.44	2.85
	633	alien	5.60	5.45	2.95
	1197	bolt	5.42	5.73	2.55
	2449	wet	5.57	5.57	3.30
	1316	cockpit	6.31	5.53	2.22
	1312	clue	5.57	5.52	2.95
	2338	thigh	5.77	5.47	2.28
	1487	endure	5.50	5.56	2.35
	2446	weight	4.28	5.40	3.27
	1754	jumble	4.41	5.30	1.23
	1326	compel	4.97	5.43	1.72
	2152	seal	5.81	5.42	2.88
	1980	plea	4.04	5.11	2.54
Negative Valence	1451	drown	2.20	6.84	2.73
	2345	threat	2.50	6.08	3.03

419	suicide	1.25	5.73	3.24
390	sickness	2.25	5.61	2.61
2113	rot	2.68	5.18	2.60
971	scandal	3.32	5.12	2.62
2240	sorrow	2.32	4.48	2.55
1427	disability	2.52	4.46	2.01
203	headache	2.02	5.07	2.86
732	dummy	3.38	4.35	2.70
1924	orphan	2.29	5.35	2.47
882	mold	3.55	4.07	2.34
906	noose	3.76	4.39	2.05
908	nuisance	3.27	4.49	2.22
86	corpse	2.18	4.74	2.71
627	ache	2.46	5.00	2.11
588	dead	1.94	5.73	4.36
1645	grieve	2.27	5.60	2.05
2264	steal	3.18	5.11	3.43
467	useless	2.13	4.87	3.01
873	messy	3.15	3.34	2.55
779	handicap	3.29	3.81	2.15
812	inferior	3.07	3.83	2.20
1858	molest	1.90	5.20	1.51

	1429	disappointment	2.37	4.60	2.44
	2473	wrong	2.93	4.67	4.43
	726	discomfort	2.19	4.17	1.91
	1981	plead	3.74	5.00	2.55
	206	helpless	2.20	5.34	2.76
	1880	nag	2.90	5.40	2.05
	964	robber	2.61	5.62	2.38
	2212	smack	3.50	5.64	2.67
	2164	sever	3.08	5.72	1.91
	392	sin	2.80	5.78	2.91
	465	upset	2.00	5.86	3.58
	1520	expel	2.77	5.90	1.85
	973	scar	3.38	4.79	2.64
	1852	miserable	1.55	5.00	3.04
	1935	pale	3.17	3.50	2.61
	720	deformed	2.41	4.07	1.83
Positive Valence	315	pillow	7.92	2.97	2.76
	183	gentle	7.31	3.21	2.93
	58	butterfly	7.17	3.47	2.45
	1323	comfortable	8.07	3.59	3.38
	118	dignified	7.10	4.12	2.69
	2294	sunflower	7.41	4.30	1.54

132	dream	6.73	4.53	3.83
355	respectful	7.22	4.60	1.99
193	grateful	7.37	4.58	3.14
1067	agree	7.00	4.59	3.52
1153	bathe	7.37	4.63	2.11
343	rainbow	8.14	4.64	2.61
268	luxury	7.88	4.75	2.49
2244	spa	7.86	4.82	2.38
1140	balloons	6.97	4.90	2.38
2462	wishful	7.50	5.57	1.73
249	knowledge	7.58	5.92	3.11
420	sunrise	7.86	5.06	2.49
175	friendly	8.43	5.11	3.12
1966	picnic	7.07	5.18	2.78
1696	honesty	7.88	5.25	2.57
145	enjoyment	7.80	5.20	1.85
1961	personality	7.48	5.88	2.91
453	trophy	7.78	5.39	2.59
502	sky	7.37	4.27	3.36
986	soft	7.12	4.63	3.21
2094	rhythm	7.06	5.73	2.75
139	elegant	7.43	4.53	2.51

2213	smart	7.93	5.90	3.69
1917	opportunity	7.41	6.47	3.32
2154	security	7.28	4.22	3.68
293	nature	7.65	4.37	3.36
404	snuggle	7.92	4.16	1.83
98	oasis	7.79	5.04	2.00
2072	relief	6.81	4.17	2.87
143	engaged	8.00	6.77	3.12
63	carefree	7.54	4.17	1.85
255	liberty	7.98	5.60	2.93
1019	treat	7.36	5.62	3.42
77	comedy	8.37	5.85	2.78

Note: Number refers to the entry number in the ANEW database (Bradley & Lang, 2010)

Appendix B: Questionnaire

[Instructions to participants:]

Congratulations! You have reached the final task of the experiment.

Please read each question thoroughly and answer as honestly and accurately as possible.

You may use the back of the paper if you run out of room.

Please notify the researcher when you are finished.

There is no time limit.

1. In the experiment, do you think you recalled some words more easily than others?

1 (a). If "yes", please explain why you think you recalled some words more easily than others.

2. Do you believe your predictions of future recall were accurate?

2 (a). Please explain your response to (2).

3. Did you notice anything about the composition of the word list(s)?

3 (a). If "yes", please explain how you think the list (s) was/were composed.

4. In general, do you believe that certain words are easier (or harder) to remember than others?

4 (a). If "yes", please explain why you think certain words vary in their memorability.

5. Did you notice any emotional words in your list(s)?

6. Do you believe emotional words are remembered differently than neutral words?

6 (a). If "yes", please explain why you believe emotional and neutral words differ in their memorability.

7. Do you believe that emotional words are remembered differently (e.g., "Happy" vs. "Sad")?

7 (a). If “yes”, please explain how you think different emotional words vary in their memorability.

8. Additionally, we would like to ask how you are feeling right now.

This scale consists of a number of words that describe difference feelings and emotions. Read each item and then mark the appropriate answer in the space next to that word. Indicate to what extent you feel this way **right now, that is, in the present moment**. Use the following scale to record your answers

1	2	3	4	5
Very slightly or not at all	A little	Moderately	Quite a bit	Extremely

_____ interested	_____ irritable
_____ distressed	_____ alert
_____ excited	_____ ashamed
_____ upset	_____ inspired
_____ strong	_____ nervous
_____ guilty	_____ determined
_____ scared	_____ attentive
_____ hostile	_____ jittery
_____ enthusiastic	_____ active
_____ proud	_____ afraid

Appendix C: Questionnaire Coding Scheme

Note: All questions (with the exception of 8) began with a yes/no question, then a short answer for part (a). 'Yes' answers were coded as 1, 'No' answers were coded as 0.

<u>Question 1 (a)</u>

1= **Primacy/Recency**: the participant mentions words at the beginning or end of the list(s). E.g. "I think I recalled words shown at the beginning of the task more easily."

2= **Word Length**: the participant states that shorter/longer words in the list(s) were easier/harder to recall. E.g. "The shorter ones were easier to remember…"

3= **Frequency**: the participant states that more common/uncommon words were easier/more difficult for them to recall. E.g. "…common words are easier."

4= **Semantic Relationships**: the participant states that they noticed some words were able to be grouped together/were similar semantically. E.g., "…some started with the same letter."

5= **Personal Significance**: the participant states that word(s) related to their own lives are easier to recall. E.g., "Some words I could relate to my personal life."

6= **Emotion**: the participant explicitly mentions that word(s) that are more positive/negative/emotional are more memorable. E.g. "Words of emotion, both negative and positive ones, were easier to recall for me."

7= **Other:** the participant mentions something that does not easily fit into a cognitive framework, or mentions something indistinguishable. E.g. "Some words worked together well or related to each other…others seemed completely unrelated and were easier to recall."

Question 2 (a)

1= **Confidence in Accuracy**: the participant clearly states they knew they recalled the words they gave higher ratings to (or vice versa) [they answered “yes” to the first part of the question]. E.g. “I think my predictions were as accurate as I could have got them.”

2= **No recall confidence**: the participant states they couldn’t remember the words they thought they would. E.g. “No, I recalled less words than I thought.”

3= **Inattention**: the participant states they may have been rehearsing the words during this time and did not provide accurate responses, or they may explicitly state they were not attentive. E.g. “I used this time to remember the words as best I could and didn’t focus on my prediction.”

4= **Reevaluation**: the participant states that they realized their ratings were inaccurate at a later point in the experiment, or during the questionnaire, once they gave it more thought. E.g. “After doing other tasks it was hard to remember the words.”

5= **Inadvertent Retrieval**: the participant states that their pattern of recall was opposite to their predictions. E.g. “Things I thought I would remember I did not remember, but things I thought I wouldn’t remember, I remembered.”

6= **Unclear:** the participant does not answer the question correctly, or their reasoning is unclear. E.g. “I think that most people will forget the words after only seeing them once and them having to do another task and then try and remember them again.”

7= **Overconfidence:** the participant states that they were initially overconfident in their judgements, and did not think they recalled as much as they thought they would. E.g. “I couldn’t remember all the ones I rated above 5”.

Question 3 (a)

1= **Word Length**: the participant discusses words specific to how short or long they were. E.g. "…longer words harder to spell."

2= **Frequency**: the participant states that the word list(s) had words used in everyday conversations (or vice versa). E.g. "It seemed that some of the words were used commonly in daily language…"

3= **Semantic Relationships**: the participant states that they noticed some words were able to be grouped together/were similar semantically. E.g. "Some of the words were often related for example the word stressful and annoyance…"

4= **Positive Emotion**: the participant explicitly states that they noticed many "happy" words. E.g. "The second list was happier…easier to recall."

5= **Negative Emotion**: the participant explicitly states that they noticed many "sad" words. E.g. "The first list of words included some that were associated with negative emotions…"

6= **All Valence**: the participant noticed that one word list was negative and neutral and the other was positive and neutral, or they mention "emotion" in general (not specific to code 4 or 5). E.g. "In list A there were pleasant words mixed in with more mundane words whereas in list B there were unpleasant words mixed in."

Question 4(a)

1= **Primacy/Recency**: the participant specifically discusses words at the beginning or end of the list(s).

2= **Word Length**: the participant discusses words specific to how short or long they were. E.g. "…longer words harder to spell."

3= **Frequency**: the participant states that the word list(s) had words used in everyday conversations (or vice versa). E.g. "It seemed that some of the words were used commonly in daily language…"

4= **Semantic Relationships**: the participant states that they noticed some words were able to be grouped together/were similar semantically. E.g. "Some of the words were often related for example the word stressful and annoyance…"

5= **Personal Significance**: the participant states that word(s) related to their own lives are easier to recall. E.g. "Some words make you think of past experiences."

6= **Emotion**: the participant explicitly mentions that word(s) that are more positive/negative/emotional are more memorable. E.g. "Words that create emotion and are more emotional words so they are easier to remember."

7= **No belief:** the participant did not answer because they responded "no" to the first part of the question.

8= **Other:** the participant mentions something that does not easily fit into a cognitive framework, or does not explicitly answer the question. E.g. "Words like 'assume' have ways to remember how to spell them…"

<u>Question 6 (a)</u>

1= **Relates to a memory**: the participant states that emotional words are more memorable because they have significance to their life, they provide an example of a memory, or they explicitly state that emotional words "remind" them of the past. E.g. "The emotional words more easily associate with previous memory, whereas you likely have no personal connection with neutral words".

2= **Relates to a current emotion**: the participant states that emotional words are more memorable because you "experience" them, or they state that reading them evokes the feeling represented by the word in the moment. E.g. "If they play on your emotions and affect your emotions in some way, you will probably remember them more than ones that had no effect on you."

3= **Stronger word meaning**: the participant discusses the quality of the word itself, in comparison to non-emotional words. E.g. "Emotional words often have a stronger meaning behind them and neutral words come across as irrelevant."

4= **Neural mechanisms**: the participant provides some neurological explanation (regardless of whether it has merit). E.g. "Memory centers in the brain are extremely close to/overlap with emotional centers…"

5= **Memory technique:** the participant states emotional words are used to employ a memory technique with, such as imagability or categorical organization. E.g. "…they make you picture something, whereas neutral words do not."

6= **No answer**: the participant has responded "no" to the previous part of the question.

7= **Other**: the participant does not explicitly answer the question, or relates it to their own performance in the experiment. E.g. "Because I could mostly only remember emotional words and the neutral words weren't as memorable."

<u>Question 7(a)</u>

1= **Variance determined by subject**: the participant states that individual differences plays a role in the memorability of different emotional words. E.g. "I think that happy words are more easily remembered in optimistic people and sad words are more easily remembered in pessimistic people…"

2= **Vary due to relation to a memory:** the participant states that it depends on how well the words relate to something in memory. E.g. "Different emotional words vary in their memorability because you can associate them with memories/encounters in your own life."

3= **Vary due to characteristics of the words**: the participant mentions something about the "strength" or arousal of words. E.g. "I think the powerful emotional words were the most memorable…"

4= **Vary due to differences in the environment**: the participant states there are more instances of positive/negative events in the world/their life and this determines the influence of emotional words. E.g. "Negative words are more likely to be remembered because there are more happy things in the world than sad"

5= **No answer:** the participant has responded "no" to the previous part of the question.

6= **Other:** the participant did not correctly address the question. E.g. "Because sad is more of a negative emotion, a higher tendency to remember it."

7= **Neural Mechanisms:** the participant provides a neurological explanation. E.g. "…they may evoke a different feeling and hence may be stored in memory differently on the basis of the feeling that that word evokes."

Table C1 *Count of responses across questions in the post-test questionnaire of Experiment 1*

		Code Number								
		0	1	2	3	4	5	6	7	8
Question										
1		2	64							
	1(a)		6	9	9	11	23	19	11	
2		46	20							
	2(a)		20	12	1	9	18		7	
3		22	44							
	3(a)		2	3	10	17	15	11	22	4
4		2	64							
	4(a)			10	24	5	29	17	1	7
5		5	61							
6		6	60							
	6(a)		31	20	6	1	3	7	4	
7		28	38							
	7(a)		12	11	8	5	28	6	1	

Note: Total percentage per question may add up to > 100 as multiple codes per participant were possible.

Table C2 *Count of responses across questions in the post-test questionnaire of experiment 2*

	Code Number								
	0	1	2	3	4	5	6	7	8
Question									
1	5	85							
1(a)		5	10	17	19	40	16	13	
2	60	30							
2(a)		30	11	0	8	25	5	24	
3	33	57							
3(a)		3	6	30	1	2	17	1	4
4	7	83							
4(a)		2	16	28	7	40	17	7	10
5	16	74							
6	7	83							
6(a)		46	26	16	5	6	7	6	
7	32	58							
7(a)		24	15	14	8	32	7	2	

Note: Total percentage per question may add up to > 100 as multiple codes per participant were possible.