10-1-2017

Post-learning Arousal Enhances Veridical Memory And Reduces False Memory In The Deese-Roediger-McDermott Paradigm

Kristy A. Nielson
Marquette University, kristy.nielson@marquette.edu

Anthony N. Correro
Marquette University

Post-learning Arousal Enhances Veridical Memory And Reduces False Memory In The Deese-Roediger-McDermott Paradigm

Kristy A. Nielson
Department of Psychology, Marquette University,
Department of Neurology and the Center for Imaging Research,
Medical College of Wisconsin,
Milwaukee, WI

Anthony N. Correro II
Department of Psychology, Marquette University,
Milwaukee, WI

Abstract: The Deese-Roediger-McDermott (DRM) paradigm examines false memory by introducing words associated with a non-presented 'critical lure' as memoranda, which typically causes the lures to be remembered as frequently as studied words. Our prior work has shown enhanced veridical memory and reduced misinformation effects when arousal is induced *after* learning (i.e., during memory consolidation). These effects have not been examined in the DRM task, or with signal detection analysis, which can elucidate the mechanisms underlying memory alterations. Thus, 130 subjects

studied and then immediately recalled six DRM lists, one after another, and then watched a 3-min arousing ($n = 61$) or neutral ($n = 69$) video. Recognition tested 70 min later showed that arousal induced after learning led to better delayed discrimination of studied words from (a) critical lures, and (b) other non-presented 'weak associates.' Furthermore, arousal reduced liberal response bias (i.e., the tendency toward accepting dubious information) for studied words relative to all foils, including critical lures and 'weak associates.' Thus, arousal induced after learning effectively increased the distinction between signal and noise by enhancing access to verbatim information and reducing endorsement of dubious information. These findings provide important insights into the cognitive mechanisms by which arousal modulates early memory consolidation processes.

Keywords: Arousal, Activation-monitoring theory, False memory, Consolidation, Signal detection theory

1. Introduction

Memories are not snapshots of events that individuals have experienced (Bartlett, 1932). In fact, some things are better remembered than others, while some recollections are distorted or inaccurate. The Deese-Roediger-McDermott (DRM; Deese, 1959; Roediger & McDermott, 1995) paradigm demonstrates this distinction, providing an experimental framework to elicit and assess the activation of false memory traces while also assessing veridical memory. In this task, lists of related words constructed around a specific, but non-presented theme word (i.e., "critical lure"), are studied. The critical lures are frequently misremembered on a later recognition test as having been previously studied (Deese, 1959; Roediger & McDermott, 1995).

The degree to which words on a list are associated with a critical lure, also known as their associative strength, is particularly important to how the dominant theories explain the DRM phenomenon. The activation-monitoring theory and the fuzzy-trace theory, each a "dual process" theory of false memory, make similar predictions regarding how the DRM paradigm is able to foment false memories. In activation-monitoring theory (Roediger, Watson, McDermott, & Gallo, 2001), it is proposed that studying the associates causes spreading activation that ultimately activates the lure within semantic networks. This leads to a source monitoring error during testing whereby individuals believe that the critical lure was studied when it is actually

new. Activation-monitoring theory is supported by findings of increased false memory when there is greater associative strength within word lists, and findings that DRM lists produce more false memory than categorized lists, where gist is high, but associative strength is lower (see Gallo, 2010). With fuzzy-trace theory, it is proposed that information is encoded into two distinct memory traces: one verbatim and one gist (e.g., Brainerd, Reyna, & Kneer, 1995). Verbatim traces contain item-specific information representing the subjective experience of encoding, while gist traces represent common semantic attributes amongst the encoded words. The critical lures from DRM themes are strong gist traces that readily create false memories. Fuzzy trace theory is supported by false memory occurring for perceptually similar but unfamiliar or unknown pictures and objects that have no pre-existing semantic associations, and through the presence of unrelated items in recognition testing yielding enhanced false memory compared with lists only including semantic associates (see Gallo, 2010). Thus, both activation-monitoring theory and fuzzy-trace theory can explain false memories through the semantic relatedness of list items to critical lures.

The DRM task has also been used to study the effects of emotion and arousal on memory. Currently, there are discrepancies in the literature regarding the directionality of the effect and what actually causes it. An important distinction between emotion and arousal helps elucidate some of the current controversies. Emotion typically refers to an affective state experienced by a person in response to a particular object or situation (Clore & Huntsinger, 2007), which can be positive or negative (i.e., valence), while arousal is defined as a range of vigilance, alertness, sympathetic activation, or responsiveness to stimuli (Revelle & Loftus, 1992; Russell, 1980). Importantly, emotion and arousal are not orthogonal; both are components of an emotional experience (Revelle & Loftus, 1992), and it is difficult to identify stimuli that produce highly positive or negative valence ratings without also producing high arousal ratings (Lang, Greenwald, Bradley, & Hamm, 1993).

Studies that focused generally on the effect of valence have differed somewhat from studies that attempted to disentangle valence from arousal in the DRM. For example, Storbeck and Clore (2005) examined emotion irrespective of arousal by eliciting positive or

negative mood using music during encoding. Negative mood led to a decrease in false memory, but positive mood and no mood induction resulted in a similar rate of false recall (although greater than the negative mood condition). The results were explained using an affect-as-information framework (Clore et al., 2001), suggesting that negative affect triggers item-specific processing (i.e., a focus on individual list items leading to verbatim memory traces), while positive affect elicits relational processing (i.e., prominently gist-based memories are formed). In contrast, studies using stimuli that span both valence and arousal dimensions have often pointed to the importance of arousal to DRM effects. One such study examined each dimension using mood induction for valence (i.e., elated, depressed, or neutral) and exercise for arousal. While mood induction failed to influence memory, low arousal led to an "overly general" style of autobiographical memory retrieval (McBride & Cappeliez, 2004). Van Damme (2013) attempted to expand upon such work in a series of three experiments. Her study used the DRM paradigm with one of six pre-learning, music-based mood conditions spanning the valence-arousal spectrum: control (no mood induction), neutral, negative/high arousal (angry), negative/low arousal (sad), positive/high arousal (happy), and positive/low arousal (serene). Each list was followed by an arithmetic distraction task. Retention of all lists en masse was assessed immediately after the final list/distraction procedure. The results showed a reduction of false memory in the higher arousal conditions, regardless of valence (see Anderson, Wais, & Gabrieli, 2006; English & Nielson, 2010; Nielson & Powless, 2007). Furthermore, signal detection analyses revealed that the arousal effect was due to enhanced discriminability of studied words versus critical lures and reduced liberal responding to critical lures.

Most studies to date have manipulated emotion or arousal before or during learning (e.g., Buchanan & Lovallo, 2001; Corson & Verrier, 2007; Kensinger & Corkin, 2004; Mather et al., 2006; Storbeck & Clore, 2005; Van Damme, 2013). In such studies, the effects of the manipulation may be on any aspect of the memory process, including attention, encoding, motivation, rehearsal, and consolidation; the effect on a specific aspect of the memory process cannot be isolated. Thus, differences in timing or context of the manipulation may contribute to some of the discrepancies in the literature. Memory consolidation, first proposed by Müller and Pilzecker

(1900), is the foundational process of assuring memory storage for later retrieval. It consists of a complex series of neurobiological processes that occur over time after the original learning (see McGaugh, 1990, 2000, 2013, 2015; Nielson & Powless, 2007; Revelle & Loftus, 1992; Torras-Garcia, Portell-Cortes, Costa-Miserachs, & Morgado-Bernal, 1997). Thus, many animal studies and some human studies have isolated the memory consolidation phase by manipulating arousal *after* encoding (e.g., Anderson et al., 2006; English & Nielson, 2010; Nielson & Arentsen, 2012; Nielson & Jensen, 1994; Nielson & Lorber, 2009; Nielson & Powless, 2007; Nielson, Yee, & Erickson, 2005). These studies, as well as others (e.g., Wang, 2013), have consistently demonstrated that moderate arousal, regardless of direction of mood or valence, enhances delayed memory. The retention interval for these studies has ranged from roughly 30 min to one week after learning, with remarkably comparable effects. Indeed, Anderson et al. (2006) showed that such effects are demonstrable after a few minutes.

Arousal as a modulator of memory consolidation is practical from an evolutionary standpoint. That is, by enhancing retention of arousing information and events, arousal allows an organism to remember and distinguish between important and mundane experiences (McGaugh, 1990). Multiple endogenous substances that are released during moderately arousing situations have been linked with the modulation of memory consolidation. These typically result from the adrenergic response to arousal and include epinephrine, norepinephrine, and glucose (see Czech, Nielson, & Laubmeier, 2000; Gold & McCarty, 1981; Nielson, Czech, & Laubmeier, 1999; Nielson & Jensen, 1994; van Stegeren, Everaerd, Cahill, McGaugh, & Gooren, 1998). Through multi-faceted pathways, adrenergic substances influence amygdala activity, which in turn influences hippocampal neurons, followed by broader systems changes in cortical and other subcortical regions involved in memory (e.g., Clewett, Sakaki, Nielsen, Petzinger, & Mather, 2017; Mather, Clewett, Sakaki, & Harley, 2015; McGaugh, 2006; McGaugh & Roozendaal, 2002). The effects of each of these substances showing the typical inverted-U dose-response patterns such that moderate doses tend to enhance memory while small or large doses can impair it (McGaugh, 2000; Yerkes & Dodson, 1908). Indeed, higher levels of arousal, such as situations involving fear or threat, are more likely to be characterized as stressful and

produce a corticosteroid response, which can impair memory retrieval processes and exacerbate false memory (Pardilla-Delgado, Alger, Cunningham, Kinealy, & Payne, 2016; Payne, Nadel, Allen, Thomas, & Jacobs, 2002; Smeets, Jelicic, & Merckelbach, 2006).

As the process of consolidation takes time, the retention interval may limit or allow for the phase and extent of consolidation to be examined. Earlier stage consolidation processes (often termed cellular consolidation), which are thought of as primarily hippocampal processes with secondary cortical contributions, are more likely tapped with retention tests occurring within hours of learning (see Genzel & Wixted, 2017). As more hours, days or weeks pass, which also allows for sleep to occur, systems consolidation, which taps more extensive and cortical contributions to memory consolidation, is more likely to have occurred (see Genzel & Wixted, 2017). The retention interval for human studies that manipulate arousal to influence memory consolidation has ranged from roughly minutes to one or two weeks after learning, thereby traversing these early and later stages. Yet, the studies with delays ranging from over 30 min to weeks have produced remarkably comparable retention profiles (e.g., Anderson et al., 2006; English & Nielson, 2010; Nielson & Arentsen, 2012; Nielson & Jensen, 1994; Nielson & Lorber, 2009; Nielson & Powless, 2007; Nielson et al., 2005). Thus, retention tests occurring within hours of learning may only reflect early or shorter term consolidation processes and it is uncertain, without longer delayed testing, whether those effects will last. The consistency in findings across such studies suggests that they may invoke sufficient initiation of consolidation processes to directly or indirectly engage longer term processes that lead to lasting memory enhancement (e.g., Mather et al., 2015). For example, when previous schematic knowledge is available, systems consolidation can much more rapidly, employing both hippocampal and systems mechanisms such as the prefrontal cortex (e.g., Coutanche & Thompson-Schill, 2014, 2015; Cowansage et al., 2014; Tse et al., 2007, 2011; van Kesteren, Fernandez, Norris, & Hermans, 2010; van Kesteren, Ruiter, Fernandez, & Henson, 2012). Stimuli such as arousal also are known to selectively modulate amygdala activity, which leads to modulation of hippocampal and other cortical and subcortical memory processes (Clewett et al., 2017; Mather et al., 2015; McGaugh, 2004, 2015). This selective modulation is believed to effectively prioritize or 'tag' memoranda, raising its salience (or 'gain'), and thereby improving its

competitive advantage in memory against lower priority information (Mather & Sutherland, 2011). Thus, early stage processes of consolidation can result in or facilitate long-term retention.

The current study attempted to partially replicate and extend our prior studies of arousal induced memory enhancement, as well as extending the work of Van Damme (2013), by isolating the arousal manipulation to the early consolidation phase in the DRM paradigm. Participants studied DRM lists and afterward were presented with either a negatively arousing or neutral video of moderate intensity. Delayed recognition was assessed 70 min later. The recognition test consisted of studied items, critical lures, unrelated unstudied foils, and weaker, unstudied associate items. The weak associates were a unique addition to the study, designed to examine the effects of arousal on more weakly associated "false" items. We hypothesized that arousal induced after learning would facilitate veridical recognition (e.g., Nielson & Powless, 2007) while also reducing false recognition of both critical lures and weak associates. Also following Van Damme (2013), we hypothesized that these effects would be due to enhanced discriminability and a less liberal response bias for studied items versus critical lures. Although weak associates were expected to be retrieved less frequently than critical lures due to their weaker associative strength with list items (see McEvoy, Nelson, & Komatsu, 1999), they were still expected to be activated at encoding (Hicks & Hancock, 2002; McEvoy et al., 1999; Roediger et al., 2001) and thereby produce effects similar to critical lures.

2. Materials and methods

2.1. Participants

Participants (N = 130; 90 female, 40 male) were adult undergraduate students who received course credit (M_{age} = 19.48, SD = 1.29). Participant demographic information is shown in Table 1. All procedures were reviewed and approved by Marquette University's Institutional Review Board.

Table 1. Group descriptive statistics (M (SD)) by group prior to experimental manipulation.

	Neutral (n = 69)	Arousal (n = 61)	p
Age (years)	19.48 (1.1)	19.48 (1.5)	0.99
Self-reported GPA	3.21 (0.5)	3.29 (0.4)	0.35
Sex	19 Male, 50 Female	21 Male, 40 Female	0.40
Recall (prior to manipulation)			
Studied words (of 90)	50.90 (7.9)	51.86 (6.9)	0.45
Critical lures (of 6)	2.95 (1.5)	2.74 (1.5)	0.41
Intrusions	1.22 (1.1)	1.31 (1.3)	0.67
Baseline subjective ratings			
Mood	5.64 (1.5)	5.70 (1.1)	0.78
Arousal	3.70 (1.7)	3.95 (1.7)	0.40

GPA = grade point average (4-point scale); p represents t tests except for sex (χ^2); Mood scale 1 (extremely negative) to 10 (extremely positive); Arousal scale 1 (not at all aroused) to 10 (extremely aroused).

2.2. Materials

2.2.1. DRM

Six DRM word lists, chosen from those that produced high rates of critical lure endorsement in Deese's (1959) original experiment (Roediger & McDermott, 1995), were compiled from normative data (Stadler, Roediger, & McDermott, 1999) and recorded by a female experimenter presenting lists at a rate of one word every two seconds (see Appendix A). Each word list, organized around a critical lure that was not presented during the encoding phase, included the 15 associates most likely to elicit the critical lure. The 6 lists were counterbalanced into 6 different orders such that each individual list occurred in each serial position. For each group session, only 1 of the 6 orders was presented. The list for the session was chosen pseudo-randomly.

2.2.2. Emotional rating scale

Subjective mood and emotional arousal were assessed using the Emotion Rating Scale (ERS; Nielson & Powless, 2007). This scale requires participants to rate their current mood on a scale of 1, *extremely negative,* to 10, *extremely positive*. Separately, this scale also asks that participants label their current arousal level on a scale of

1, *not at all aroused,* to 10, *extremely aroused*. This measure is simpler and independent, yet comparable to the simultaneous measures in the grid-based instrument used by Van Damme (2013) to measure mood and arousal.

2.2.3. Arousal manipulation

Emotional arousal was manipulated using one of two videos ($N_{arousal}$ = 61, $N_{neutral}$ = 69). Participants in odd-numbered experimental sessions watched the arousal video, which was a 3-min clip of live-action oral surgery. This clip has been shown to elicit moderate subjective emotional arousal and physiological arousal in prior studies (English & Nielson, 2010; Nielson & Arentsen, 2012; Nielson & Powless, 2007; Nielson et al., 2005; Stone & Nielson, 2001). Participants in even-numbered experimental sessions watched the neutral video, which was a 3-min clip from a PBS documentary concerning the link between heart disease and depression. Prior studies have indicated that this clip is interesting enough to maintain attention without substantively raising arousal level or changing mood (English & Nielson, 2010; Nielson & Arentsen, 2012). The videos did not overlap semantically with each other or the DRM lists. After viewing and after rating their own mood and arousal state (ERS), participants were asked to rate the clip they viewed (scale of 1, "not at all" to 10, "extremely") on four dimensions: unpleasant, funny, disgusting and interesting.

2.2.4. Delayed retention

Six 'brain teaser' problems and 14 questionnaires assessing various opinions and attitudes were presented as filler during a 70-min delay interval (not analyzed for the present study). The delayed recognition test employed the classic Remember/Know paradigm (Gardiner, 1988; Tulving, 1985), where a "Remember" response indicated the participant could distinctly remember the item, a "Know" response indicated that the participant knew the word was presented but could not remember anything specific about its presentation, and a "No" response indicated the item was not presented during the encoding phase. The recognition test consisted of 162 items presented in pseudo-random order: 90 previously studied list items (all 15 from each list), the 6 previously un-presented critical lures, 12 previously

un-presented 'weak associates' of the studied lists (2 per DRM list), and 54 new, unrelated items (i.e., "foils"). The foils were highly imageable, concrete nouns taken from established norms (Paivio, Yuille, & Madigan, 1968). Weak associates (see Appendix A) were words that we determined were associated to the superordinate DRM lists based on the following criteria: each appeared as a synonym of its associated critical lure in an English-language thesaurus; none appeared in the top 15 associates used in DRM lists (Stadler et al., 1999); and none appeared as a frequently associated word with its critical lure or other DRM list items in word association databases (Russell & Jenkins, 1954; Toglia & Battig, 1978).[1]

2.3. Procedure

Experimental sessions were conducted in a group format over one 120-min session. Informed consent was obtained at the beginning of each session, followed by a demographic survey. The DRM lists were then presented one at a time, with instructions to remember the words, followed by immediate free recall after each list, as is common in most list-learning experiments to ensure attention and encoding. This was used despite the potential for increased false memory effects (Roediger & McDermott, 1995) as the primary interest was to investigate arousal effects on false and veridical memory and both experimental groups would experience the same conditions. After all lists were completed, the first ERS was obtained as a pre-manipulation baseline. Next, participants watched either the oral surgery video (arousal group) or the documentary (neutral group). This was followed by a second ERS assessment to document the direct influence of the manipulation. The 'brain teasers' and various surveys were then administered until 70 min had elapsed. Recognition testing was administered following this delay. Finally, participants were debriefed and dismissed.

2.3.1. Analysis approach

All analyses were conducted using SPSS Version 21.0. A significance value of $p < 0.05$ was used for all statistical tests. Group comparability and manipulation checks were conducted using t-tests and mixed ANOVA. Memory analyses were conducted using MANCOVA

(with one-way ANCOVA follow-up analyses) in order to examine the entire set of dependent measures for the effects of arousal and to examine whether sex differences contributed to the results. Although the manipulation groups were comparable in sex distribution (despite a heavy bias toward female participants), sex was covaried because men and women tend to differ in performance on verbal memory tasks (Halpern, 2000; Herlitz, Nilsson, & Bäckman, 1997; Kimura, 1999; Ruff, Light, & Quayhagen, 1989) and some prior studies have reported sex differences in memory after arousal manipulations (Cahill, 2003; Cahill et al., 2001; Canli, Desmond, Zhao, & Gabrieli, 2002).

Hit and false alarm rates reflect both sensitivity and response bias, without differentiation (Stanislaw & Todorov, 1999). Sensitivity refers to the degree of overlap between the signal and noise distributions, while response bias is the general tendency to respond either "yes" or "No." This is typically reported using the parametric statistic d' ($= z$ (false alarms) $- z$ (hits)), which is used herein for studied words; larger values indicate greater sensitivity. To better distinguish these contributions to recognition performance, additional signal detection analyses were used. Non-parametric indices have been the preferred signal detection analysis metrics for DRM studies (e.g., Van Damme, 2013), particularly when examining false memory (i.e., critical lures). As such, sensitivity and response bias were examined using A' and B'', respectively. A' ranges from 0 to 1. A value of 0.5 indicates chance performance, and larger values indicate greater sensitivity. B'' ranges from -1 to 1, where negative values indicate the tendency to accept dubious information as correct (i.e., liberal responding) and positive values reflect the tendency to reject dubious information as incorrect (i.e., conservative responding). A post hoc analysis of arousal effect on critical target endorsement dependent on initial recall was conducted using X^2.

3. Results

3.1. Demographic data and manipulation checks

Descriptive statistics for the demographic characteristics and manipulation checks can be found in Table 1. There were no significant differences between the groups with respect to age, self-reported

grade-point average, or sex distribution. Immediate free recall performance, measured after each list, was examined to assure that the groups appropriately attended to the encoding stimuli. As expected, the groups did not differ in the number of list items recalled, $t(128) = -0.76$, $p = 0.45$, $d = 0.13$, in the number of critical lures falsely recalled, $t(128) = 0.84$, $p = 0.41$, $d = 0.15$, or in the number of intrusions, $t(128) = -0.43$, $p = 0.67$, $d = 0.08$. Similarly, the two groups were comparable, prior to manipulation, in baseline mood ratings, $t(128) = -0.29$, $p = 0.78$, $d = 0.05$, and arousal ratings, $t(128) = -0.86$, $p = 0.39$, $d = 0.15$.

To demonstrate the effect of the arousal manipulation, change in mood and arousal ratings was evaluated using 2 (group) × 2 (measures) mixed ANOVA. For mood, the arousal manipulation (i.e., oral surgery video) led to more negative mood after manipulation in the arousal group. This was evident in a significant main effect of group, $F(1, 128) = 4.35$, $p = 0.04$, $\eta_p^2 = 0.03$, a significant main effect of measures, $F(1, 128) = 32.79$, $p < 0.001$, $\eta_p^2 = 0.20$. These were qualified by a significant interaction, $F(1, 128) = 12.58$, $p = 0.001$, $\eta_p^2 = 0.09$. Follow up contrasts showed a significant reduction in the arousal group ($p < 0.001$) but no significant change in the neutral group ($p = 0.78$); see Fig. 1. Similarly for arousal ratings, the manipulation produced a significantly greater change, in this case an increase, in the arousal group. This was evidenced by significant main effects of arousal, $F(1, 128) = 11.65$, $p = 0.001$, $\eta_p^2 = 0.08$) and measures, $F(1, 128) = 75.02$, $p < 0.001$, $\eta_p^2 = 0.37$, which were qualified by a significant interaction, $F(1, 128) = 30.30$, $p < 0.001$, $\eta_p^2 = 0.13$. Follow up contrasts showed a significant increase in the arousal group ($p < 0.001$) but no significant change in the neutral group ($p = 0.39$); see Fig. 1.
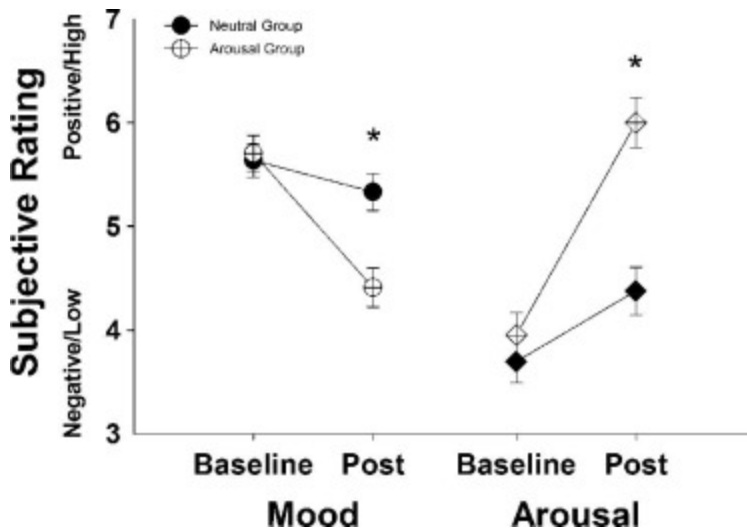
Fig. 1. Mood and arousal ratings shown by group (scale = 1–10, negative to positive (mood) or low to high (arousal)), at baseline (prior to manipulation) and post-manipulation. Mood was made significantly more negative and arousal rating significantly increased in the arousal group (n = 61) relative to the neutral group (n = 69) showing the manipulation was effective.

Ratings of the videos further showed the distinction in effect of the manipulation. The arousal video (oral surgery) was rated as significantly more unpleasant ($p < 0.001$; mean$_{arousal}$ = 6.9, SD = 2.6; mean$_{neutral}$ = 4.1, SD = 2.0) and disgusting ($p < 0.001$; mean$_{arousal}$ = 6.5, SD = 2.7; mean$_{neutral}$ = 1.4, SD = 1.0) than the neutral video, while the neutral video (documentary) was rated as significantly more interesting ($p = 0.001$; mean$_{arousal}$ = 5.1, SD = 2.5; mean$_{neutral}$ = 6.4, SD = 1.9). The neutral video was also 'funnier' than the arousal video ($p = 0.014$), but ratings were near floor in both groups (mean$_{arousal}$ = 1.3, SD = 0.9; mean$_{neutral}$ = 1.8, SD = 1.4; scale = 1–10 where 1 = 'not at all funny').

3.2. Recognition analyses

3.2.1. MANCOVA

All delayed memory metrics were first submitted as dependent variables in a MANCOVA using arousal condition as the independent variable and sex as the covariate (see Methods). The dependent variables included raw measures and signal detection metrics for studied words, critical lures, and 'weak associates.' The overall model was significant for arousal condition, $F(11, 117) = 2.21$, $p = 0.018$,

η_p^2 = 0.172, but sex was not a significant covariate overall, $F(11, 117)$ = 1.28, p = 0.246, η_p^2 = 0.107. What follows is the univariate ANCOVA for each metric. Although sex was not significant overall, it was significant for two metrics and as such, the approach was retained through the analyses.

3.2.2. Raw responses

The raw response metrics are shown for each group in Fig. 2. The hit rate for studied words did not significantly differ by group. For "Remember" responses, the mean hit rate was nearly identical between the groups at 71%, $F(1, 127)$ = 0.00, p = 0.991, η_p^2 = 0.000; it was also similar for "Know" responses, which were far less frequent at 11–13%, $F(1, 127)$ = 0.48, p = 0.491, η_p^2 = 0.004. Total hits also failed to differ by group, $F(1, 127)$ = 1.10, p = 0.297, η_p^2 = 0.009, but it covaried significantly with sex (p = 0.04, females > males; a full factorial model showed no interaction of sex with arousal condition, p > 0.60).
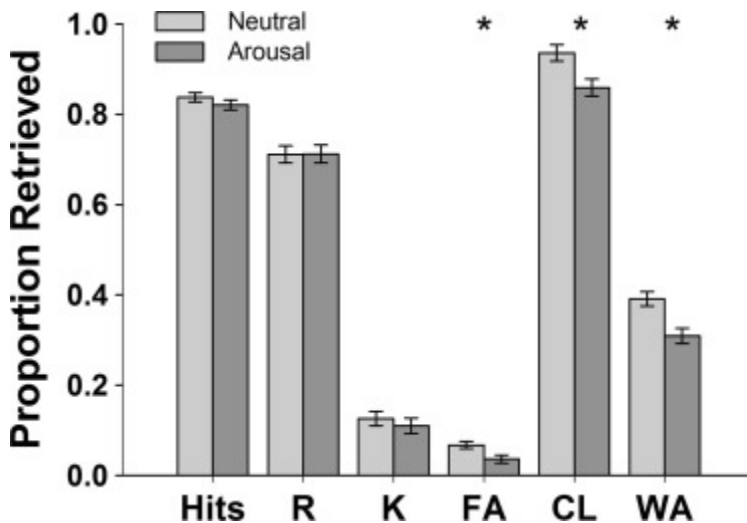

Fig. 2. Results are shown (mean (±SEM)) for each raw score memory metric by arousal condition (arousal group n = 61; neutral group n = 69). Hits = total recognition hits (i.e., previously presented words correctly endorsed); R = recognition hits marked "Remember"; K = recognition hits marked "Know"; FA = False Alarms (i.e., foils incorrectly endorsed as having been presented); CL = critical lures endorsed (i.e., false memory); WA = weak associates endorsed (i.e., false memory). Arousal led to reduced false alarms and reduced false memory.

Importantly, in contrast to hits, the false alarm rate of unstudied foils was significantly lower in the arousal group vs. the neutral group, $F(1, 127) = 6.59$, $p = 0.01$, $\eta_p^2 = 0.05$, as was the rate of endorsing either critical lures as "old" (i.e., as previously presented), $F(1, 127) = 8.48$, $p = 0.004$, $\eta_p^2 = 0.06$, or 'weak associates' as "old", $F(1, 127) = 12.05$, $p = 0.001$, $\eta_p^2 = 0.09$. Overall, the false alarm rate was also greater for critical lures than for 'weak associates,' as hypothesized based on the task design, $t(1\ 2\ 9) = 35.33$, $p < 0.001$.

To further clarify and contextualize the raw indices of word retrieval and group differences, as well as to allow better comparisons with other recognition memory studies, traditional d' measures of recognition accuracy were also computed. d' was significantly different between the arousal groups overall, $F(1, 127) = 3.94$, $p = 0.049$, $\eta_p^2 = 0.030$; mean$_{arousal} = 3.41$, SEM = 0.13; mean$_{neutral} = 3.06$, SEM = 0.12. As the vast majority of the responses were Remember responses, the result was comparable between arousal groups for Remember responses, $F(1, 127) = 3.86$, $p = 0.052$, $\eta_p^2 = 0.030$, but not significantly different for the more rare Know responses ($F(1, 127) = 2.11$, $p = 0.149$, $\eta_p^2 = 0.016$). Taken together, these results indicate that arousal after learning led to better veridical memory and reduced false memory.

3.2.3. Signal detection analyses

The signal detection metrics are shown for each group in Fig. 3. Sensitivity (A') showed no difference by arousal group for studied words relative to unrelated foils, $F(1, 127) = 1.66$, $p = 0.199$, $\eta_p^2 = 0.013$. This was the only other metric for which sex was a significant covariate ($p = 0.04$, females > males). However, arousal after learning led to better discrimination of studied items from critical lures, $F(1, 127) = 4.37$, $p = 0.039$, $\eta_p^2 = 0.033$, and studied items from 'weak associates', $F(1, 127) = 7.97$, $p = 0.006$, $\eta_p^2 = 0.059$. Thus, although sensitivity to studied words versus unrelated foils was

not enhanced by arousal, the ability to discriminate dubious lures and 'weak associates' from studied items was improved by arousal.
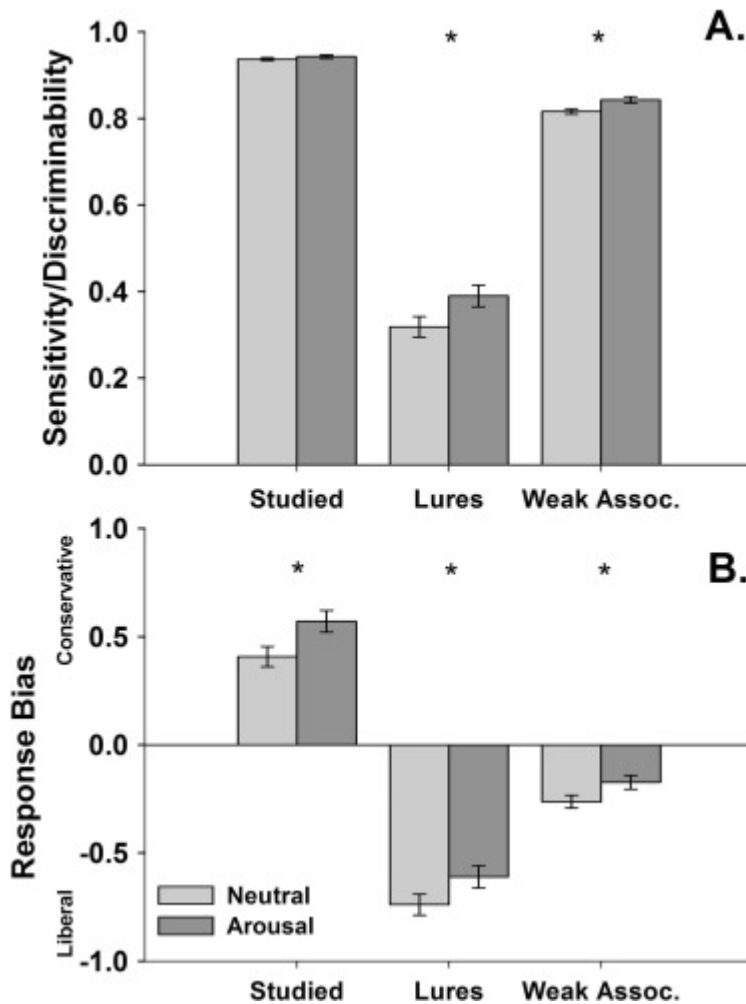

Fig. 3. Results are shown (mean (±SEM)) for signal detection metrics by arousal condition (arousal group n = 61; neutral group n = 69). Panel A = Sensitivity/discriminability (*A'*, range = 0–1); Panel B = Response bias (*B"*, range = −1 to 1*)*; Studied = studied words (hits) relative to unrelated foils; Lures = critical lures relative to hits; Weak Assoc = weak associates relative to hits. *A'* demonstrated greater discriminability for studied versus false information (i.e., critical targets and weak associates) in the arousal condition. *B"* demonstrated more conservative responding to studied words relative to unrelated foils (i.e., better veridical memory) and less liberal responding to dubious words (i.e., less false memory) in the arousal condition.

Response bias (*B"*) to studied words versus unrelated foils was conservative, and arousal induced after learning resulted in an increase of that conservative bias, $F(1, 127) = 5.62$, $p = 0.02$, $\eta_p^2 = 0.042$. In contrast, the response bias comparing critical lures to

studied words was quite liberal (i.e., a tendency to accept the dubious lure item relative to actual studied items). Arousal after learning trended toward reducing that liberal bias, $F(1, 127) = 3.23$, $p = 0.075$, $\eta_p^2 = 0.025$. Further supporting this trend, the response bias comparing 'weak associates' to studied words was somewhat liberal, and arousal induced after learning significantly reduced that liberal tendency, $F(1, 127) = 4.15$, $p = 0.044$, $\eta_p^2 = 0.032$. Thus, arousal participants were even more conservative about foil words, relative to studied items, than neutral participants. In addition, the tendency to accept dubious lures and 'weak associates' as actually having been presented was significantly reduced by arousal.

3.2.4. Post-hoc analysis of critical targets

Cross tabulation of critical target (i.e., lure) recall and recognition by arousal group is shown in Table 2 with χ^2 statistics. The nature of the DRM data suggest caution in interpreting these data as there are a number of cells with 0 observations (also precluding valid $2 \times 2 \times 2$ log-linear analysis). The results suggest that when critical targets were erroneously recalled immediately after learning, there was nearly 100% chance of later endorsing them as having been studied. As there is a ceiling effect, it is difficult to draw strong conclusions, but it appeared that arousal had a negligible effect on strongly activated critical lures (arousal = 97% endorsement after being recalled; neutral = 99% endorsement after being recalled). In contrast, examination of critical lures that were not initially recalled showed that participants in the arousal group had a significantly reduced likelihood of erroneously endorsing them during delayed retention testing (77% vs. 88%).

Table 2. Cross-tabulation of critical lure endorsement by group and immediate recall.

		Recognition endorsement												
		Anger		Chair		Rough		Needle		Sleep		Sweet		All targets
Group	Recall	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Neutral	No	2	26	4	35	0	31	5	23	10	36	3	32	24	183
(n = 69)	Yes	0	41	0	30	0	38	0	41	2	21	1	33	3	204
$\chi^2 p$		0.16		0.13		N/A		0.01		0.31		0.32		0.09	
Cramer V		0.21		0.22				0.34		0.16		0.12		0.35	

		Recognition endorsement													
		Anger		Chair		Rough		Needle		Sleep		Sweet		All targets	
Group	Recall	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
OR		0.13		0.13				0.05		0.34		0.32		0.23	
Arousal	No	5	25	10	23	0	28	12	22	9	26	11	29	47	153
(n = 61)	Yes	0	31	2	26	0	33	0	27	1	25	1	20	4	162
$\chi^2 p$		**0.02**		**0.03**		N/A		**0.00**		**0.03**		**0.03**		**0.05**	
Cramer V		0.30		0.29				0.44		0.29		0.27		0.38	
OR		0.07		0.18				0.03		0.12		0.13		0.12	

Recall was tested before the arousal manipulation; Chi-Squared p value is Fisher's exact (bold values are p ≤ .05); Cramer V = effect size; OR = odds ratio, odds of endorsing a critical target as 'studied' if it was not recalled. For example, if not recalled, the lure "anger" was 87% less likely to be endorsed during recognition in the Neutral group, and 93% less like to be endorsed during recognition in the Arousal group.

4. Discussion

The purpose of the current study was to evaluate whether a brief, moderately arousing video shown after learning can alter early memory consolidation processes to influence both veridical and false memory. The DRM paradigm was used because it is ideally suited to investigate both veridical and false retrieval in the same task (Roediger & McDermott, 1995). We hypothesized that post-learning arousal would enhance the number of studied items accurately recognized while also reducing the number of critical lures endorsed. This reduction in false recognition was proposed to occur through enhanced sensitivity of studied items versus critical lures and more conservative responding elicited by arousal (Van Damme, 2013).

The results supported the hypotheses. Arousal induced after learning resulted in better discrimination of studied items and reduced false retrieval of foils, critical lures and other non-presented, 'weak associates.' Although overall critical lure retrieval was reduced by arousal, examining the match between recall and later recognition of lures clarified that the effect of arousal on lures depended on whether they were recalled. Specifically, false memory for lures was reduced by

arousal if they were not recalled immediately after learning. However, lures that were recalled produced consistent false memory in both arousal groups, indicating that once recalled, these items effectively acted as studied items and were influenced by arousal in a comparable manner.

The results of the current study are consistent with our earlier study using a different task, the misinformation paradigm, but the same post-learning manipulations and a longer retention interval (English & Nielson, 2010). Although arousal did not directly enhance recognition hit rate in the present study, it enhanced veridical memory by reducing the rate of endorsing unrelated 'foils' as having been studied (i.e., false alarms). This is evidenced by the significant group difference in d' and the raw false alarm measure. In addition, B'' showed that by increasing resistance to erroneous information through more conservative responding, arousal increased accuracy. Moreover, arousal induced after learning improved veridical memory principally through reduced retrieval of critical lures via more conservative responding (B'') to the lures and weak associates. It furthermore increased sensitivity (A') to studied words relative to the critical lures. These results importantly replicate the primary results of Van Damme (2013) despite the differences in approach (including post- versus pre-learning arousal induction and the presence versus absence of immediate recall testing). Although any stage of memory could be influenced by arousal induced during or before learning, the comparability of current post-learning results with those of Van Damme (2013) suggests that arousal effects on early consolidation may be responsible for memory enhancement in both pre- and post-learning induction designs.

A novel aspect of the current study was the inclusion of secondary, unstudied lures on the recognition test. These 'weak associates' were comparable to other studied list items in their association with the critical lure, but like the critical lure, they were not studied. These items allowed the investigation of the effects of arousal on weaker false items than the primary theme word itself, while also giving added opportunities to examine false memory; the DRM typically affords only one false memory opportunity per list. We expected that the 'weak associates' would be activated during encoding but at a lower rate than the critical lures. Arousal was further

expected to reduce false alarms for the 'weak associates' through more conservative responding. The results were consistent with these hypotheses. Comparable to the findings for critical lures, this reduction in false memory was due to greater sensitivity to studied items versus 'weak associates' and a more conservative response bias in the arousal group.

Our results could be interpreted as consistent with either of the dominant dual framework models of memory. Fuzzy-trace theory proposes that item-specific information is stored in verbatim traces, while semantically related information occurs in gist traces that can cause false retrieval (Brainerd et al., 1995). Our results would imply that verbatim traces (which would include studied words and recalled critical lures) were more distinctive or easily activated at test as a result of arousal induced after learning, while gist traces were weaker or suppressed by arousal. The effect of arousal on response bias toward a more conservative, error-resistant approach is also fitting with the proposition that lower arousal state results in an over-general retrieval style, while higher arousal results in greater item-specificity (McBride & Cappeliez, 2004). Alternatively activation-monitoring theory proposes that the DRM triggers spreading semantic activation, which results in false perception-like qualities of related but unstudied items, generating false memory due to source monitoring errors. This is particularly notable when the items have high associative strength (Gallo, 2010; Roediger et al., 2001), as they did in the present study. By this interpretation, arousal appeared to combat spreading activation or access to related but false information at test, while improving access to studied information. This is consistent with studies using post-learning arousal with other tasks that also showed specific enhancement of accurate recollection and reduction of familiarity-based or false retrieval (e.g., Anderson et al., 2006; English & Nielson, 2010). The present results also align with a study in which immediate post-learning arousal selectively enhanced exam performance for the material in a professor's lecture relative to other learned information from the course, both related and unrelated, in prior and subsequent lectures for the course (Nielson & Arentsen, 2012). By either dual framework model account, post-learning arousal appeared to enhance 'signal' and reduce 'noise', which has been purported as the role of arousal in memory in many animal and human studies (see McGaugh, 2000). The exception to this interpretation is related to critical lures

that were recalled. The recall of these lures created very strong source confusion, which was not reduced by arousal.

A recent theory is perhaps also helpful in interpreting the present results. Arousal Biased Competition (ABC) theory posits that arousal modulates competition for mental representation on the basis of top-down goals or perceptual salience, resulting in enhanced memory for high priority information and reduced memory for low priority information (Mather & Sutherland, 2011). In a word-list learning task, arousal induced after learning would enhance memory for the words as they were the focus of attention and thus, prioritized, while it would reduce memory for words not receiving attention or receiving less attention. Our results are thus consistent with what ABC theory would predict. Given the semantic relatedness of the DRM words, the finding of reduced false alarms to unrelated foils reinforces that arousal reduced the priority of items unrelated to the studied words. More, the DRM paradigm presents a compelling test of ABC theory in that study of the words (prioritized items) provokes semantic activation of related but non-presented words. The current methods used lists with high rates of association with critical lures and thus, lure endorsement (Roediger & McDermott, 1995). Thus, it is expected that each lure provoked strong semantic activation. However, some lures were recalled during immediate testing, while some were not. The act of recall appeared to enhance priority, making these lures comparable to studied words, as they became the focus of attention during recall (Mather & Sutherland, 2011). In contrast, lures that were not recalled were semantically activated (as evidenced by a high rate of endorsement), yet they remained low priority and thus, arousal reduced their later retrieval. Taken with the signal detection metrics, arousal participants had better sensitivity to studied information and more resistance to dubious information, so long as that dubious information did not become prioritized through recall.

Extensive literature in animal models and humans demonstrates that for memory modulation to occur peripheral and central responses to arousal must result in amygdala activity, which selectively modulates hippocampal and other cortical and subcortical memory processes, such as noradrenergic and glutamate release and long-term potentiation (Clewett et al., 2017; Mather et al., 2015; McGaugh, 2004, 2015). The enhancement of veridical memory and reduction of

false memory in the present study suggests that arousal modulated amygdala input differently for consolidation of studied items (i.e., high priority or 'signal') versus contextually relevant but lower priority, unstudied items. The former likely occurred via medial temporal/hippocampal activity while the latter would have been associated with the medial prefrontal cortex, the retrosplenial cortex, and late visual processing in the occipital cortex (Aminoff, Schacter, & Bar, 2008; Kim & Cabeza, 2007). Just as there is ample evidence indicating that amygdala selectively prioritizes emotional stimuli over neutral stimuli when they occur together, our results indicate that amygdala stimulation during consolidation can assist in distinguishing high and low priority from amongst neutral stimuli even when those stimuli are difficult to distinguish due to high semantic relatedness. Yet, if priority is established (whether via study or false recall), amygdala activity is likely to enhance it. While the mechanisms underlying these effects have not been investigated in the DRM, Mather et al. (2015) propose that glutamate-noradrenaline amygdalar loops ultimately influence synaptic plasticity to mediate gain for storage of prioritized memoranda.

Arousal can alter various stages of the memory process, including attentional focus and consolidation (see Christianson & Loftus, 1990; Eysenck, 1976; Heuer & Reisberg, 1990; McGaugh, 2006; Nielson, Radtke, & Jensen, 1996; Revelle & Loftus, 1992). Post-learning manipulations isolate the effects of arousal on consolidation, which are typically evident in greater effects after a long versus a short retention delay. Specifically, very short retention intervals typically evidence nil or negative effects of moderate arousal on retrieval while longer delays typically reveal enhancing effects (see Eysenck, 1976; Heuer & Reisberg, 1990; Kleinsmith & Kaplan, 1963; LaBar & Phelps, 1998; Nielson et al., 2005; Revelle & Loftus, 1992). The present study used a 70-min retention interval in order to allow for sufficient early consolidation processes to occur (Anderson et al., 2006; McGaugh, 1966; Nielson & Powless, 2007; Nielson et al., 1996, 2005), but this does not discount that further consolidation continues afterward. Memory consolidation is considered to have overlapping shorter- and longer-term phases with both contributing to long-lasting memory (McGaugh, 2000). The processes examined in experiments like the current one may only address shorter-term processes, while those associated with long-lasting memory enhancement may require

retention tests that are delayed 24 h or longer and likely involve additional long-term processes of both consolidation and reconsolidation (Dudai, 2004; McGaugh, 2000). Yet, a number of prior studies using similar post-learning paradigms with retention delays from 24 h to weeks (e.g., English & Nielson, 2010; Nielson & Arentsen, 2012; Nielson & Bryant, 2005; Nielson & Jensen, 1994; Nielson & Meltzer, 2009; Nielson & Powless, 2007; Nielson et al., 2005) have produced extremely comparable effects and effect sizes. For example, the degree of improvement in memory performance due to arousal in these past studies is quite similar to the current results at approximately 10%. This degree of improvement was notably shown to be enough to aid older adults to the level of unaided young adult memory performance (Nielson & Jensen, 1994). Thus, although retention intervals of days or weeks more confidently represent durable memory enhancement, the comparability of the current findings with those of prior studies using much longer retention delays suggests that the current paradigm engaged early stage consolidation processes that could result in or facilitate lasting effects (see Genzel & Wixted, 2017). Indeed, when previous schematic knowledge is available (which the DRM was designed to exploit) or when particularly salient stimulation occurs, systems consolidation can occur, engaging both hippocampal and cortical networks (e.g., prefrontal cortex), more rapidly than is typically seen (e.g., Coutanche & Thompson-Schill, 2014, 2015; Cowansage et al., 2014; Tse et al., 2007, 2011; van Kesteren et al., 2010, 2012). As noted by Mather and Sutherland (2011), arousal prioritizes stimuli as particularly salient and thus, this manipulation may have sufficiently engaged early-stage processes to lead to long-term retention, or it may have rapidly engaged both short and long-term consolidation processes. Although these issues were not the focus of the current study, future systematic studies with short and long retention intervals are warranted to clarify these processes and specifically address whether the results in the current study are likely to represent lasting modulation effects.

Only a negatively valenced arousal stimulus was investigated in this study, which subjectively increased arousal and negatively altered mood. This manipulation was contrasted by a neutral stimulus that did not significantly alter arousal or mood ratings. Although it is impossible to definitively assert that arousal rather than mood, or valence, was responsible for the present results, arousal is implicated by the larger

literature. First, both positive and negative emotional stimuli that evoke subjective arousal also typically produce an effect on valence ratings (e.g., Lang et al., 1993, Figs. 1 and 2). Indeed, this is one reason that the animal literature has often employed exogenous administration of arousal-associated hormones (see McGaugh, 2000) and some human studies employ non-emotive, sympathetic arousal-inducing manipulations such as muscle tension (Nielson & Jensen, 1994; Nielson, Wulff, & Arentsen, 2014; Nielson et al., 1996) and adrenergic-inhibiting agents such as beta-blockers to study these mechanisms (Cahill, Prins, Weber, & McGaugh, 1994; Nielson & Jensen, 1994). Furthermore, studies that have attempted to distinguish arousal and valence effects in human memory modulation studies using emotive stimuli have implicated arousal as the dimension of primary relevance (Anderson et al., 2006; Corson & Verrier, 2007; McBride & Cappeliez, 2004; Nielson & Powless, 2007; Schwartz, 1975; Van Damme, 2013). Lastly, the arousal stimulus used in the present study has been used in past studies where arousal, as measured by a change in heart rate and/or skin conductance, was specifically responsible for the effects on memory (e.g., Nielson & Meltzer, 2009; Nielson et al., 2005). This was particularly clear in a study of alexithymia, where participants exhibited memory enhancement by arousal and associated physiological responses to arousal, but they did not describe the stimulus as subjectively arousing (Nielson & Meltzer, 2009). Thus, while mood or valence effects cannot be ruled out, the preponderance of the evidence implicates arousal, not valence, as responsible for memory modulation.

The current study employed recall testing following the study of each list. This procedure is typical in list-learning experiments, to assure that attention and encoding occur. Yet, this procedure has been shown to increase false memory effects (e.g., Roediger & McDermott, 1995). Importantly, studied word and lure retrieval in the present study was consistent with, albeit at lower proportions, than in Van Damme's (2013, Exp. 1) study after which we modeled it. The difference is likely due to the longer retention delay in the current study. Thus, while false memory may have been increased by the inclusion of an immediate retrieval procedure, both experimental groups experienced the same conditions, and the procedure neither raised false memory levels higher than in typical studies nor created ceiling effects.

No arousal or mood assessments were made prior to learning or directly prior to the recognition test. Instead, baseline was taken immediately prior to manipulation and the effect of manipulation was measured directly afterward. One could argue that arousal had not dissipated prior to retention testing and thus, the state of the groups could have differed at testing, thereby producing the effects we report. This is, however, highly unlikely. Acute adrenergic responses dissipate within minutes (see Robertson, 2004). Furthermore, we have shown that the subjective effects of similar emotional arousal stimuli resolve within a few minutes (Nielson & Lorber, 2009). Even the slower cortisol response (i.e., stress), which is likely negligible with this type of task (Lovallo, Farag, Vincent, Thomas, & Wilson, 2006), is expected to peak within 10–30 min and resolve within the 70-min delay interval we imposed (Dickerson & Kemeny, 2004; Kirschbaum, Strasburger, & Langkrar, 1993). Finally, a study with an one-week retention interval very similarly showed that post-learning arousal, using the same two videos in the present study, enhanced accurate retrieval and reduced false memory in a "misinformation" task (English & Nielson, 2010). Thus, taken together, these studies strongly suggest that the memory enhancement effects in the current study were not attributable to lingering arousal state at retrieval.

The present study demonstrated that arousal induced after learning DRM lists enhanced memory performance for studied information and reduced false memory on a 70-min delayed retention test. We demonstrated similar effects using a misinformation paradigm in an earlier study (English & Nielson, 2010) after a one-week delay. Furthermore, the current findings showed that arousal altered memory through enhancing the sensitivity of studied items relative to lures and resistance to dubious information through more conservative responding. Thus, these results support and extend the literature indicating that arousal induced after learning enhances early memory consolidation processes leading to greater distinction between 'signal' and 'noise' and add to our understanding of the manner in which this cognitive distinctiveness is produced. Taken with other recent evidence, there is indication that the distinctiveness arousal affords to memoranda can be robust, sustained, and may have utility in everyday life (e.g., Nielson & Arentsen, 2012).

Acknowledgements

Appendix A

Studied words with non-presented critical targets in header (bold):

Anger	**Chair**	**Rough**	**Needle**	**Sleep**	**Sweet**
Mad	Table	Smooth	Thread	Bed	Sour
Fear	Sit	Bumpy	Pin	Rest	Candy
Hate	Legs	Road	Eye	Awake	Sugar
Rage	Seat	Tough	Sewing	Tired	Bitter
Temper	Couch	Sandpaper	Sharp	Dream	Good
Fury	Desk	Jagged	Point	Wake	Taste
Ire	Recliner	Ready	Prick	Snooze	Tooth
Wrath	Sofa	Coarse	Thimble	Blanket	Nice
Happy	Wood	Uneven	Haystack	Doze	Honey
Fight	Cushion	Riders	Thorn	Slumber	Soda
Hatred	Swivel	Rugged	Hurt	Snore	Chocolate
Mean	Stool	Sand	Injection	Nap	Heart
Calm	Sitting	Boards	Syringe	Peace	Cake
Emotion	Rocking	Ground	Cloth	Yawn	Tart
Enrage	Bench	Gravel	Knitting	Drowsy	Pie

Weak associates:

Hostile	Beach	Rocky	Hypodermic	Lullaby	Syrup
Annoyed	Bleacher	Chapped	Stitch	Hibernate	Sticky

References

Aminoff et al., 2008. E. Aminoff, D.L. Schacter, M. Bar. The cortical underpinnings of context-based memory distortion. *Journal of Cognitive Neuroscience*, 20 (12) (2008), pp. 2226-2237, 10.1162/jocn.2008.20156

Anderson et al., 2006. A.K. Anderson, P.E. Wais, J.D. Gabrieli. Emotion enhances remembrance of neutral events past. *Proceedings of the National Academy of Sciences of the United States of America*, 103 (5) (2006), pp. 1599-1604, 10.1073/pnas.0506308103

Bartlett, 1932. F.C. Bartlett. *Remembering: A study in experimental and social psychology.* Macmillan, New York, NY (1932)

Brainerd et al., 1995. C.J. Brainerd, V.F. Reyna, R. Kneer. False-recognition reversal: When similarity is distinctive. *Journal of Memory and Language*, 34 (2) (1995), pp. 157-185, 10.1006/jmla.1995.1008

Buchanan and Lovallo, 2001. T.W. Buchanan, W.R. Lovallo. Enhanced memory for emtional material following stress-level cortisol treatment in humans. *Psychoneuroendocrinology*, 26 (2001), pp. 307-317, 10.1016/S0306-4530(00)00058-5

Cahill, 2003. L. Cahill. Sex- and hemisphere-related influences on the neurobiology of emotionally influenced memory. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 27 (8) (2003), pp. 1235-1241, 10.1016/j.pnpbp.2003.09.019

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

Cahill et al., 1994. L. Cahill, B. Prins, M. Weber, J.L. McGaugh. Beta-adrenergic activation and memory for emotional events. *Nature*, 371 (1994), pp. 702-704

Canli et al., 2002. T. Canli, J. Desmond, Z. Zhao, J.D.E. Gabrieli. Sex differences in the neural basis of emotional memories. *Proceedings of the National Academy of Sciences USA*, 99 (2002), pp. 10789-10794, 10.1073/pnas.162356599

Christianson and Loftus, 1990. S.Å. Christianson, E.F. Loftus. Some characteristics of people's traumatic memories. *Bulletin of the Psychonomic Society*, 28 (3) (1990), pp. 195-198

Clewett et al., 2017. D. Clewett, M. Sakaki, S. Nielsen, G. Petzinger, M. Mather. Noradrenergic mechanisms of arousal's bidirectional effects on episodic memory. *Neurobiology of Learning and Memory*, 137 (2017), pp. 1-14, 10.1016/j.nlm.2016.10.017

Clore and Huntsinger, 2007. G.L. Clore, J.R. Huntsinger. How emotions inform judgment and regulate thought. *Trends in Cognitive Sciences*, 11 (9) (2007), pp. 393-399, 10.1016/j.tics.2007.08.005

Clore et al., 2001. G.L. Clore, R.S. Wyer, B. Dienes, K. Gasper, C. Gohm, L. Isbell. Affective feelings as feedback: Some cognitive consequences. L.L. Martin, G.L. Clore (Eds.), *Theories of mood and cognition: A user's guidebook*, Erlbaum, Mahwah, NJ (2001), pp. 27-62

Corson and Verrier, 2007. Y. Corson, N. Verrier. Emotions and false memories: Valence or arousal? *Psychological Science*, 18 (3) (2007), pp. 208-211, 10.1111/j.1467-9280.2007.01874.x

Coutanche and Thompson-Schill, 2014. M.N. Coutanche, S.L. Thompson-Schill. Fast mapping rapidly integrates information into existing

memory networks. *Journal of Experimental Psychology: General*, 143 (6) (2014), pp. 2296-2303, 10.1037/xge0000020

Coutanche and Thompson-Schill, 2015. M.N. Coutanche, S.L. Thompson-Schill. Rapid consolidation of new knowledge in adulthood via fast mapping. *Trends in Cognitive Sciences*, 19 (9) (2015), pp. 486-488, 10.1016/j.tics.2015.06.001

Cowansage et al., 2014. K.K. Cowansage, T. Shuman, B.C. Dillingham, A. Chang, P. Golshani, M. Mayford. Direct reactivation of a coherent neocortical memory of context. *Neuron*, 84 (2) (2014), pp. 432-441, 10.1016/j.neuron.2014.09.022

Czech et al., 2000. D.A. Czech, K.A. Nielson, K.K. Laubmeier. Chronic administration of propranolol impairs Morris water maze retention in the mouse. *Neurobiology of Learning and Memory*, 74 (2000), pp. 17-26, 10.1006/nlme.1999.3944

Deese, 1959. J. Deese. On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, 58 (1) (1959), pp. 17-22, 10.1037/h0046671

Dickerson and Kemeny, 2004. S.S. Dickerson, M.E. Kemeny. Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin*, 130 (3) (2004), pp. 355-391

Dudai, 2004. Y. Dudai. The neurobiology of consolidations, or, how stable is the engram. *Annual Review of Psychology*, 55 (2004), pp. 51-89, 10.1146/annurev.psych.55.090902.142050Feb

English and Nielson, 2010. S.M. English, K.A. Nielson. Reduction of the misinformation effect by arousal induced after learning. *Cognition*, 117 (2) (2010), pp. 237-242, 10.1016/j.cognition.2010.08.014

Eysenck, 1976. M.W. Eysenck. Arousal, learning, and memory. *Psychological Bulletin*, 83 (3) (1976), pp. 389-404

Gallo, 2010. D.A. Gallo. False memories and fantastic beliefs: 15 years of the DRM illusion. *Memory and Cognition*, 38 (7) (2010), pp. 833-848, 10.3758/MC.38.7.833

Gardiner, 1988. J.M. Gardiner. Functional aspects of recollective experience. *Memory and Cognition*, 16 (4) (1988), pp. 309-313, 10.3758/BF03197041

Genzel and Wixted, 2017. L. Genzel, J.T. Wixted. Cellular and systems consolidation of declarative memory. N. Axmacher, B. Rasch (Eds.), *Cognitive neuroscience of memory consolidation*, Springer, Switzerland (2017), pp. 3-16

Gold and McCarty, 1981. P.E. Gold, R. McCarty. Plasma catecholamines: Changes after footshock and seizure-producing frontal cortex stimulation. *Behavioral and Neural Biology*, 31 (1981), pp. 247-260, 10.1016/S0163-1047(81)91259-0

Halpern, 2000. D.F. Halpern. Sex differences in cognitive abilities. (3rd ed.), Lawrence Erlbaum Associates Inc., Mahwah, NJ (2000)

Herlitz et al., 1997. A. Herlitz, L.-G. Nilsson, L.G. Bäckman. Gender differences in episodic memory. *Memory and Cognition*, 25 (1997), pp. 801-811, 10.3758/BF03211324

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

Hicks and Hancock, 2002. J.L. Hicks, T.W. Hancock. Backward associative strength determines source attributions given to false memories. *Psychonomic Bulletin Review*, 9 (4) (2002), pp. 807-815, 10.3758/BF03196339

Kensinger and Corkin, 2004. E.A. Kensinger, S. Corkin. Two routes to emotional memory: Distinct neural processes for valence and arousal. *Proceedings of the National Academy of Sciences of the United States of America*, 101 (9) (2004), pp. 3310-3315, 10.1073/pnas.0306408101

Kim and Cabeza, 2007. H. Kim, R. Cabeza. Differential contributions of prefrontal, medial temporal, and sensory-perceptual regions to true and false memory formation. *Cerebral Cortex*, 17 (9) (2007), pp. 2143-2150, 10.1093/cercor/bhl122

Kimura, 1999. D. Kimura. *Sex and cognition.* MIT Press, Cambridge, MA (1999)

Kirschbaum et al., 1993. C. Kirschbaum, C.J. Strasburger, J. Langkrar. Attenuated cortisol response to psychological stress but not to CRH or ergometry in young habitual smokers. *Pharmacology, Biochemistry and Behavior*, 44 (3) (1993), pp. 527-531

Kleinsmith and Kaplan, 1963. L.J. Kleinsmith, S. Kaplan. Paired-associate learning as a function of arousal and interpolated interval. *Journal of Experimental Psychology*, 65 (2) (1963), pp. 190-193

LaBar and Phelps, 1998. K.S. LaBar, E.A. Phelps. Arousal-mediated memory consolidation: Role of the medial temporal lobe in humans. *Psychological Science*, 9 (6) (1998), pp. 490-493

Lang et al., 1993. P.J. Lang, M.K. Greenwald, M.M. Bradley, A.O. Hamm. Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, 30 (3) (1993), pp. 261-273

Lovallo et al., 2006. W.R. Lovallo, N.H. Farag, A.S. Vincent, T.L. Thomas, M.F. Wilson. Cortisol responses to mental stress, exercise, and meals following caffeine intake in men and women. *Pharmacology, Biochemistry and Behavior*, 83 (3) (2006), pp. 441-447, 10.1016/j.pbb.2006.03.005

Mather et al., 2015. M. Mather, D. Clewett, M. Sakaki, C.W. Harley. Norepinephrine ignites local hot spots of neuronal excitation: How

arousal amplifies selectivity in perception and memory. *Behavioral and Brain Sciences* (2015), pp. 1-100, 10.1017/S0140525X15000667

Mather et al., 2006. M. Mather, K.J. Mitchell, C.L. Raye, D.L. Novak, E.J. Greene, M.K. Johnson. Emotional arousal can impair feature binding in working memory. *Journal of Cognitive Neuroscience*, 18 (4) (2006), pp. 614-625, 10.1162/jocn.2006.18.4.614

Mather and Sutherland, 2011. M. Mather, M.R. Sutherland. Arousal-biased competition in perception and memory. *Perspectives on Psychological Science*, 6 (2) (2011), pp. 114-133, 10.1177/1745691611400234

McBride and Cappeliez, 2004. C. McBride, P. Cappeliez. Effects of manipulating valence and arousal components of mood on specificity of autobiographical memory. *Psychological Reports*, 95 (2) (2004), pp. 615-630, 10.2466/pr0.95.2.615-630

McEvoy et al., 1999. C.L. McEvoy, D.L. Nelson, T. Komatsu. What is the connection between true and false memories? The differential roles of interitem associations in recall and recognition. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 25 (5) (1999), pp. 1177-1194, 10.1037/0278-7393.25.5.1177

McGaugh, 1966. J.L. McGaugh. Time-dependent processes in memory storage. *Science*, 153 (3742) (1966), pp. 1351-1358

McGaugh, 1990. J.L. McGaugh. Significance and remembrance: The role of neuromodulatory systems. *Psychological Science*, 1 (1990), pp. 15-25, 10.1111/j.1467-9280.1990.tb00060.x

McGaugh, 2000. J.L. McGaugh. Memory – a century of consolidation. *Science*, 287 (2000), pp. 248-251, 10.1126/science.287.5451.248 (14 January)

McGaugh, 2004. J.L. McGaugh. The amygdala modulates the consolidation of memories of emotionally arousing experiences. *Annual Review of Neuroscience*, 27 (2004) (2004), pp. 1-28, 10.1146/annurev.neuro.27.070203.144157

McGaugh, 2006. J.L. McGaugh. Make mild moments memorable: Add a little arousal. *Trends in Cognitive Sciences*, 10 (8) (2006), pp. 345-347

McGaugh, 2013. J.L. McGaugh. Making lasting memories: Remembering the significant. *Proceedings of the National Academy of Sciences of the United States of America*, 110 (Suppl 2) (2013), pp. 10402-10407, 10.1073/pnas.1301209110

McGaugh, 2015. J.L. McGaugh. Consolidating memories. *Annual Review of Psychology*, 66 (2015), pp. 1-24, 10.1146/annurev-psych-010814-014954

McGaugh and Roozendaal, 2002. J.L. McGaugh, B. Roozendaal. Role of adrenal stress hormones in forming lasting memories in the brain. *Current Opinion in Neurobiology*, 12 (2) (2002), pp. 205-210

Müller and Pilzecker, 1900. G.E. Müller, A. Pilzecker. Experimentelle beitrage zur lehre vom gedachtnis [Experimental contributions in learning and

memory]. *Zeitschrift für Psychologie Erganzungsband*, 1 (1900), pp. 1-288

Nielson and Arentsen, 2012. K.A. Nielson, T.J. Arentsen. Memory modulation in the classroom: Selective enhancement of college examination performance by arousal induced after lecture. *Neurobiology of Learning and Memory*, 98 (1) (2012), pp. 12-16, 10.1016/j.nlm.2012.04.002

Nielson and Bryant, 2005. K.A. Nielson, T. Bryant. The effects of non-contingent extrinsic and intrinsic rewards on memory consolidation. *Neurobiology of Learning and Memory*, 84 (2005), pp. 42-48

Nielson et al., 1999. K.A. Nielson, D.A. Czech, K.K. Laubmeier. Chronic administration of propranolol impairs inhibitory avoidance retention in the mouse. *Neurobiology of Learning and Memory*, 71 (1999), pp. 248-257, 10.1006/nlme.1998.3873

Nielson and Jensen, 1994. K.A. Nielson, R.A. Jensen. Beta-adrenergic receptor antagonist antihypertensive medications impair arousal-induced modulation of working memory in elderly humans. *Behavioral and Neural Biology*, 62 (1994), pp. 190-200, 10.1016/S0163-1047(05)80017-2

Nielson and Lorber, 2009. K.A. Nielson, W. Lorber. Enhanced post-learning memory consolidation is influenced by arousal predisposition and emotion regulation but not by stimulus valence or arousal. *Neurobiology of Learning and Memory*, 92 (2009), pp. 70-79, 10.1016/j.nln.2009.03.002

Nielson and Meltzer, 2009. K.A. Nielson, M.A. Meltzer. Modulation of long-term memory by arousal in alexithymia: The role of interpretation. *Consciousness and Cognition*, 18 (3) (2009), pp. 786-793, 10.1016/j.concog.2009.06.001

Nielson and Powless, 2007. K.A. Nielson, M. Powless. Positive and negative sources of emotional arousal enhance long-term word-list retention when induced as long as thirty minutes after learning. *Neurobiology of Learning and Memory*, 88 (2007), pp. 40-47, 10.1016/j.nlm.2007.03.005

Nielson et al., 1996. K.A. Nielson, R.C. Radtke, R.A. Jensen. Arousal-induced modulation of memory storage processes in humans. *Neurobiology of Learning and Memory*, 66 (1996), pp. 133-142

Nielson et al., 2014. K.A. Nielson, L.L. Wulff, T.J. Arentsen. Muscle tension induced after learning enhances long-term narrative and visual memory in healthy older adults. *Neurobiology of Learning and Memory*, 109 (2014), pp. 144-150, 10.1016/j.nlm.2014.01.008

Nielson et al., 2005. K.A. Nielson, D. Yee, K.I. Erickson. Memory enhancement by a semantically unrelated emotional arousal source

induced after learning. *Neurobiology of Learning and Memory*, 84 (2005), pp. 49-56

Paivio et al., 1968. A. Paivio, J.C. Yuille, S.A. Madigan. Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology*, 76 (Suppl.) (1968), pp. 1-25, 10.1037/h0025327

Pardilla-Delgado et al., 2016. E. Pardilla-Delgado, S.E. Alger, T.J. Cunningham, B. Kinealy, J.D. Payne. Effects of post-encoding stress on performance in the DRM false memory paradigm. *Learning and Memory*, 23 (1) (2016), pp. 46-50, 10.1101/lm.039354.115

Payne et al., 2002. J.D. Payne, L. Nadel, J.J. Allen, K.G. Thomas, W.J. Jacobs. The effects of experimentally induced stress on false recognition. *Memory*, 10 (1) (2002), pp. 1-6, 10.1080/09658210143000119

Revelle and Loftus, 1992. W. Revelle, D.A. Loftus. The implications of arousal effects for the study of affect and memory. S. Christianson (Ed.), *The handbook of emotion and memory*, Hillsdale Lawrence Erlbaum Assoc (1992), pp. 113-150

Robertson, 2004. D. Robertson (Ed.), *Primer on the autonomic nervous system* (2nd ed.), Elsevier Academic Press, San Diego, CA (2004)

Roediger and McDermott, 1995. H.L. Roediger 3rd, K.B. McDermott. Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 21 (4) (1995), pp. 8040-8814, 10.1037/0278-7393.21.4.803

Roediger et al., 2001. H.L. Roediger 3rd, J.M. Watson, K.B. McDermott, D.A. Gallo. Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin Review*, 8 (3) (2001), pp. 385-407, 10.3758/BF03196177

Ruff et al., 1989. R.M. Ruff, R.H. Light, M. Quayhagen. Selective reminding tests: A normative study of verbal learning in adults. *Journal of Clinical and Experimental Neuropsychology*, 11 (1989), pp. 539-550, 10.1080/01688638908400912

Russell, 1980. J.A. Russell. A circumplex model of affect. *Journal of Personality and Social Psychology*, 39 (6) (1980), pp. 1161-1178, 10.1037/h0077714

Russell and Jenkins, 1954. Russell, W. A., & Jenkins, J. J. (1954). *Studies on the role of language in behavior (Technical Report No. 11): The complete minnesota norms for responses to 100 words from the Kent-Rosanoff word association test.* Minneapolis, MN: University of Minnesota.

Schwartz, 1975. S. Schwartz. Individual differences in cognition: Some relationships between personality and memory. *Journal of Research in Personality*, 9 (3) (1975), pp. 217-225

Smeets et al., 2006. T. Smeets, M. Jelicic, H. Merckelbach. Stress-induced cortisol responses, sex differences, and false recollections in a DRM paradigm. *Biological Psychology*, 72 (2) (2006), pp. 164-172, 10.1016/j.biopsycho.2005.09.004

Stadler et al., 1999. M.A. Stadler, H.L. Roediger 3rd, K.B. McDermott. Norms for word lists that create false memories. *Memory and Cognition*, 27 (3) (1999), pp. 494-500, 10.3758/BF03211543

Stanislaw and Todorov, 1999. H. Stanislaw, N. Todorov. Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, 31 (1) (1999), pp. 137-149, 10.3758/BF03207704

Stone and Nielson, 2001. L.A. Stone, K.A. Nielson. Physiological response to arousal and impaired emotional expression in alexithymia. *Psychotherapy and Psychosomatics*, 70 (2001), pp. 92-102, 10.1159/000056232

Storbeck and Clore, 2005. J. Storbeck, G.L. Clore. With sadness comes accuracy; with happiness, false memory: Mood and the false memory effect. *Psychological Science*, 16 (10) (2005), pp. 785-791, 10.1111/j.1467-9280.2005.01615.x

Toglia and Battig, 1978. M.P. Toglia, W.F. Battig. Handbook of semantic word norms. Erlbaum, Hillsdale, NJ (1978)

Torras-Garcia et al., 1997. M. Torras-Garcia, I. Portell-Cortes, D. Costa-Miserachs, I. Morgado-Bernal. Long-term memory modulation by posttraining epinephrine in rats differential effects depending on the basic learning capacity. *Behavioral Neuroscience*, 111 (1997), pp. 301-308

Tse et al., 2007. D. Tse, R.F. Langston, M. Kakeyama, I. Bethus, P.A. Spooner, E.R. Wood, …, R.G. Morris. Schemas and memory consolidation. *Science*, 316 (5821) (2007), pp. 76-82, 10.1126/science.1135935

Tse et al., 2011. D. Tse, T. Takeuchi, M. Kakeyama, Y. Kajii, H. Okuno, C. Tohyama, …, R.G. Morris. Schema-dependent gene activation and memory encoding in neocortex. *Science*, 333 (6044) (2011), pp. 891-895, 10.1126/science.1205274

Tulving, 1985. E. Tulving. Memory and consciousness. *Canadian Psychologist*, 26 (1) (1985), pp. 1-12, 10.1037/h0080017

Van Damme, 2013. I. Van Damme. Mood and the DRM paradigm: An investigation of the effects of valence and arousal on false memory. *Quarterly Journal of Experimental Psychology*, 66 (6) (2013), pp. 1060-1081, 10.1080/17470218.2012.727837

van Kesteren et al., 2010. M.T. van Kesteren, G. Fernandez, D.G. Norris, E.J. Hermans. Persistent schema-dependent hippocampal-neocortical connectivity during memory encoding and postencoding rest in humans. *Proceedings of the National Academy of Sciences of the*

United States of America, 107 (16) (2010), pp. 7550-7555, 10.1073/pnas.0914892107

van Kesteren et al., 2012. M.T. van Kesteren, D.J. Ruiter, G. Fernandez, R.N. Henson. How schema and novelty augment memory formation. *Trends in Neurosciences*, 35 (4) (2012), pp. 211-219, 10.1016/j.tins.2012.02.001

van Stegeren et al., 1998. A.H. van Stegeren, W. Everaerd, L. Cahill, J.L. McGaugh, L.J.G. Gooren. Memory for emotional events: Differential effects of centrally versus peripherally acting beta-blocking agents. *Psychopharmacology* (Berl), 138 (1998), pp. 305-310, 10.1007/s002130050675

Wang, 2013. B. Wang. Effect of positive emotion on consolidation of memory for faces: The modulation of facial valence and facial gender. *Memory*, 21 (6) (2013), pp. 707-721, 10.1080/09658211.2012.753461

Yerkes and Dodson, 1908. R. Yerkes, J. Dodson. The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, 18 (1908), pp. 459-482, 10.1002/cne.920180503

[1]There was one exception to this protocol. Due to an error during experiment preparation, the word "beach" was included in place of "bench" in the studied item list for the "chair" DRM list. "Beach" was intended as a weak associate as it meets the above requirements, but the word "bench" is modestly more strongly associated with "chair;" it would normally be on the studied list. Instead, "bench" was included as a weak associate. That is, it was not studied but was included in the recognition test. "Bench" only ranked 4th highest in endorsement for weak associates (behind "rocky," "stich," and "hostile"). However, because it was possible that it might bias results due to its potentially higher association with the critical lure than intended, all analyses were also performed without this item (i.e., 11 weak associates instead of 12); no results differed even minimally through the exclusion of this item. As such, the full analysis of 12 items is presented. Appendix A shows the words in their intended lists and categories for future use.