Distribution Agreement

Signature: _____________________________ Date _______________

Experience-related eye movements and pupillary responses reflect declarative memory for emotional and neutral pictures

By

Cory S. Inman

Master of Arts

Psychology

Stephan Hamann, Ph.D.
Advisor

Patricia Bauer, Ph.D.
Committee Member

Robert Hampton, Ph.D.
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Experience-related eye movements and pupillary responses reflect declarative memory for emotional and neutral pictures

By

Cory S. Inman

B.A., Georgia State University, 2006

Advisor: Stephan Hamann, Ph.D.

An abstract of

A thesis submitted to the Faculty of the

James T. Laney School of Graduate Studies of Emory University

in partial fulfillment of the requirements for the degree of

Master of Arts

in Psychology

2010

Abstract

Experience-related eye movements and pupillary responses reflect declarative memory for emotional and neutral pictures

By Cory S. Inman

Increased arousal as elicited by an emotional stimulus is related to narrowing of attention to the stimulus. Previous studies suggest that focal attention on the salient features of an experience enhances the accuracy and vividness of memory. In fact, a previous study using eye movements as a measure of attentional allocation has shown that attention is narrower for pictures later recognized as vividly 'remembered' than just 'familiar'. However, several questions remain concerning how attentional allocation relates to emotion and subsequent memory.

The relationship between emotional valence, physiological arousal, and narrowing of attention has not been extensively examined. To examine this relationship, we measured the pupil response to emotional stimuli concurrently with eye tracking. Enhanced memory for physiologically arousing stimuli has been demonstrated, but it is unknown whether physiological arousal as measured by the pupil response is related to enhanced memory. To address these remaining questions, eye movements and pupillometry were continuously recorded while participants encoded and later retrieved positive, negative, and neutral pictures in a free recall task and a remember/know recognition task.

Pupil changes were larger when viewing emotionally arousing pictures. Additionally, vivid recollection is associated with increased pupillary responses during encoding and retrieval of emotional and neutral stimuli. Eye fixations were more clustered when encoding negative pictures that were later vividly remembered than vividly remembered positive or neutral pictures. This finding suggests that vivid recollection of a stimulus depends on how attention is allocated during encoding and the stimulus valence. Similar to previous findings, during the recognition test, eye fixations were more clustered for remembered relative to familiar pictures, especially negative remembered pictures. Finally, extending previous findings, positive stimuli were sampled more frequently than negative or neutral stimuli during encoding and retrieval. These findings suggest that vivid recollection is related to encoding of a few distinct features of highly arousing negative pictures, but for positive, low arousal negative, and neutral stimuli there is no difference between the 'remember' and 'familiar' component of recollective experience. During recognition, vivid recollection may be prompted by enhanced memory for the salient features of previously seen negative photos.

Experience-related eye movements and pupillary responses reflect declarative memory for emotional and neutral pictures

By

Cory S. Inman
B.A., Georgia State University, 2006

Advisor: Stephan Hamann, Ph.D.

A thesis submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Master of Arts
in Psychology
2010

Acknowledgments

I gratefully thank my advisor, Stephan Hamann, and my Masters committee members, Patricia Bauer and Rob Hampton, for their advice and guidance in preparation of this thesis. I also thank my lab colleagues, Jennifer Wilson, Carolina Campanella, and Katye Vytal for their thoughtful comments and discussion in preparation of this thesis. Finally, I would like to thank members of the Memory at Emory lab, Thanujeni Pathman and Ayzit Doydum for their assistance and guidance in using the eye tracking system.

Table of Contents

List of Tables and Figures

Experience-related eye movements and pupillary responses reflect declarative memory for emotional and neutral pictures

Every moment, our attention is captured by *salient* features in the immediate environment, like an *italicized,* **bold**, or emotional word in a sentence. Further attention and cognitive elaboration on the stimulus begins the creation of memories for the experience in our minds. With time, increasingly permanent traces of the original experience are stored in memory. Re-experiencing the stimulus might bring to mind a vivid recollection of the thoughts or feelings originally associated with the experience, i.e. the episodic context of the experience. This subjective feeling-of-remembering is known as the recollective experience (Tulving, 1985; Sharot & Yonelinas, 2004). In contrast, re-experiencing the stimulus might only bring to mind the feeling that the stimulus has been seen before, but no episodic context. This is known as the 'familiar' or 'know' component of recollective experience.

How attention is allocated to features of a stimulus is an important factor that affects what stimuli we remember and how we remember them (Gardiner, Gregg, and Karayianni, 2006). Previous studies suggest that focal visual attention on the salient features of an experience enhances the accuracy and vividness of memory (Rajaram, 1993; Dewhurst & Conway, 1994; Ochsner, 2000). In addition, a study using eye movements as a measure of visual attention allocation have shown that attention is narrower for pictures later recognized as vividly 'remembered' than just 'familiar' (Sharot, Davidson, Carson, & Phelps, 2008). They also found that negatively arousing pictures were related to narrowing of attention during encoding relative to neutral stimuli. These previous findings suggest that narrowing of attention during encoding is related to the emotional arousal and subsequent recollective experience of a previously seen picture.

Based on this previous study, it is still unknown whether emotional valence contributes to the narrowing of attention during encoding in addition to emotional arousal. Additionally, the relationship between physiological arousal and narrowing of attention has not been extensively examined. To examine this relationship, we measured the pupil response to emotional stimuli concurrently with eye tracking. Enhanced memory for physiologically arousing stimuli has been demonstrated by many studies, but it is unknown whether physiological arousal as measured by the pupil response is related to enhanced memory. It is also unclear whether the narrowing of attention during encoding is related to subsequent memory accuracy, i.e. whether a stimulus is later remembered versus forgotten. The current study addresses these questions by tracking visual attention allocation and measuring the pupil response during encoding and retrieval of positive, negative, and neutral stimuli. Further, we used a subsequent memory paradigm and eye tracking to assess attention allocation during retrieval of pictures that were later vividly remembered, only ‘familiar’, or forgotten.

Episodic Memory Accuracy and Recollective Experience

The ability to consciously remember events from a particular time and place is known as episodic memory. In contrast, semantic memory is our ability to remember facts about the world (Tulving, 1983). These classes of memory fall under the umbrella of declarative memory, i.e. the processes of encoding, storing, and retrieving knowledge that can be consciously and intentionally recollected

(Gabrieli, 1998). The process of encoding stimuli that were subsequently remembered or forgotten has been well studied (Haman, Ely, Grafton, & Kilts, 1999; Canli, Zhao, Brewer, Gabrieli, & Cahill, 2000; Canli, Desmond, Zhao, & Gabrieli, 2002; Mickley-Steinmetz & Kensinger, 2009; reviewed in Paller & Wagner, 2002). Less is currently known about the processes underlying storage and retrieval of stimuli, although these stages in the life of a memory are now receiving much more attention (Bauer, 2007; Buchanan, 2007).

The recollection process that operates on recognition is thought to be similar to the search processes that are utilized in tests of recall. Thus, free recall is typically used as another index of recollection, similar to the 'remember' component of recollective mmm njjjexperience (Yonelinas, 2002). An important difference between free recall tests and recognition tests is the number of cues provided during study and that match stimulus features during test. Free recall tests have no external cues to remind the participant of the original experience. In recognition tests, the match between cues during study and test is high and in some designs exactly matching.

While recall and old/new recognition tests have been very effective in answering questions about the contents and accuracy of episodic memory, they don't fully measure the subjective experience underlying the memory response. As described earlier, the autonoetic consciousness or feeling-of-remembering is the extent to which a participant can vividly recollect the episodic context in which the stimulus in question was previously experienced (Tulving, 1985). The episodic context is a thought or feeling originally associated with a previously experienced stimulus. Subsequently 'remembered' stimuli have an episodic context associated with them, while subsequently ‘known’ or ‘familiar’ stimuli lack an episodic context. Remembering is thought to involve distinctive processing that requires more attention, while familiarity is more automatic and depends on fluency (Gardiner, Gregg, and Karayianni, 2006). Examining encoding and retrieval mechanisms, like attention allocation and arousal, that lead to accurate and vivid episodic memories has implications for theories describing why we remember certain aspects of our experiences and forget others.

Attention and Emotional Memory

Attention and memory. Studies of attention and episodic memory have shown that memory accuracy increases when more attention is devoted to a stimulus (Yonelinas, 2002; Rajaram, 1993; Dewhurst & Conway, 1994; Ochsner, 2000). In addition, previous studies have shown that memory accuracy and recollective experience for features of a stimulus can be increased or decreased when attention at encoding is experimentally directed or divided. For example, when attention is intentionally devoted to features of an image, memory for those features is increased (Rajaram, 1993; Dewhurst & Conway, 1994; Ochsner, 2000). When attention is divided, recall and recognition memory accuracy has been shown to suffer, with recall accuracy suffering more than recognition accuracy (Craik, Govoni, Naveh-Benjamin, & Anderson, 1996). Subjectively, less items were vividly ‘remembered’ when attention is divided at encoding, while the effect on the familiarity of items is negligible (Yonelinas, 2001; Curran, 2004; Mangels, Picton, Craik, 2001). When attention is divided during the recognition test, less of an effect is found on the vividly remembered items (Yonelinas, 2002).

Based on this behavioral evidence for attention's affect on subjective recollective experience, Sharot and colleagues (2008) explored whether the allocation of attention resources differed as a function of recollective experiences. They proposed two possible patterns of attentional allocation during encoding and retrieval. First, remembered stimuli might be characterized by more dispersed attention patterns because subsequently 'remembered' stimuli tend to be accompanied by contextual information. Alternatively, remembered stimuli might be characterized by more focused attention on a few distinct features of the stimulus which could lead to deeper elaboration and encoding of specific salient features of the stimulus. They used emotional valence and arousal to manipulate the kind and amount of salient features in each stimulus. In particular, they showed participants negative and neutral images at encoding and subsequently tested their memory with a remember/know recognition test. During encoding and the recognition test, they measured the distance between each eye-fixation (a momentary pause in the gaze pattern) and the number of fixations per stimulus. They employed these eye-tracking methods in a subsequent memory paradigm. In the subsequent memory paradigm, memories were subsequently binned into various memory categories, like items that were correctly remembered or forgotten during encoding or categorized as remembered, familiar, or new (for review see Paller and Wagner, 2002). Then average eye-movement measurements for each these categories are obtained within the encoding and retrieval sessions.

Sharot and colleague's (2008) found that subjectively 'remembered' items were characterized by more clustered fixations relative to subjectively 'familiar' items during encoding and retrieval. Their interpretation for this finding was that attentional resources for subsequently ‘remembered’ items are focused on specific parts of a scene and thus more deeply encoded. At retrieval, they found more clustered fixations for ‘remember’ items than ‘familiar’ items and more fixations for ‘familiar’ items than ‘remembered’ items. This finding was interpreted as evidence for the idea that if focal features were deeply encoded, then less of the same features or cues were needed to produce a vivid recollective experience. Note that these effects at encoding and retrieval were regardless of the emotional valence or arousal of the stimulus, although there was also a narrowing of attention for emotional items (negative only) relative to neutral items during encoding. Other than just considering emotion for its usefulness in manipulating the salience of the stimulus, several questions particular to emotion’s influence on memory remain.

Arousal and memory. Emotion is also a potent way to enhance episodic memory (Hamann, 2001; Phelps, 2004; Labar & Cabeza, 2006). One way to characterize emotion is by a stimulus’s arousal and valence. The enhancing effects of emotional arousal on the encoding, storage, and retrieval of memories have been well tested and described in a variety of lab settings, including non-human animal studies as well as many behavioral, neuropsychological, and neuroimaging human studies (for review Squire, 2004). In the brain, these effects are partially attributed to up-regulation of amygdala activity through various endogenous hormones (McGaugh, 2000), through which bi-directional connections, influence activity in the hippocampus (Hamann, 2001). The amygdala is highly connected with other regions in the brain as well (Pessoa, 2008). In fact, the amygdala’s functional connections with the visual cortex and other sensory cortices were proposed as evidence for its influence upon visual and sensory attention (Pessoa, McKenna, Gutierrez, & Ungerleider, 2002; Pessoa, 2008). In particular, the amygdala’s functional connectivity with the fusiform gyrus (a region important for high-level visual processing) predicts the likelihood that visual details of arousing stimuli will be encoded. This connectivity between the amygdala and visual processing areas is attenuated in patients with amygdala lesions (Vuilleumier et al., 2004). The amygdala is also thought to modulate the neural response of hippocampal and visual processing regions during retrieval, perhaps through modulation of search processes (Daeselaar et al., 2007; Smith et al., 2006; Buchanan, 2007). In general, emotional arousal is thought to engage neural and cognitive processes not usually recruited unless there is an emotional reaction, in addition to driving typical encoding mechanisms like increased attention, rehearsal, and cognitive elaboration (Hamann, 2001; Kensinger, 2009).

The most well known proposal of emotional arousal’s focal influence on the allocation of attentional resources is the Easterbrook hypothesis. Easterbrook’s cue-utilization hypothesis (1959) proposed that increased emotional arousal of a stimulus serves to reduces the range and number of cues to which an organism will attend. Thus, the restriction of attentional focus may cause deeper encoding of arousing information, while failing to process other information in the stimulus (Kensinger, 2009; Hamann, 2001; Labar & Cabeza, 2006; Phelps, 2004). The cue-utilization hypothesis has been examined using a number of different methods (Calvo & Lang, 2004; Calvo & Lang, 2005), but to date Sharot et al’s (2008) study addresses the proposal most directly in human participants. In line with this hypothesis, Sharot and colleagues (2008) found the span of cues receiving attention was reduced (more clustered fixations) for emotionally negative items relative to neutral items during encoding. Inconsistent with Easterbrook’s hypothesis, they found that participants tended to fixate on more emotional items than neutral items during encoding. Taken with the findings of more clustered fixations for subjectively ‘remembered’ versus ‘familiar’ items, these findings suggest that the narrowing of attention is not unique to recollective experience or emotion, but occurs for both as a result of similar mechanisms (Sharot et al., 2008). Attention also tends to be captured by a smaller range of the salient features of emotional stimuli than neutral stimuli, but this does not necessarily lead to a higher likelihood of vivid recollection of emotional stimuli over neutral stimuli.

Attentional allocation for the central and peripheral features of a stimulus is a key encoding mechanism currently used to explain emotional arousal’s overall mnemonic benefit. The emotion-induced memory trade off (Kensinger, Garoff-Eaton, Schacter, 2007) is an effect whereby participants tend to remember the central emotional content of a scene, while forgetting the peripheral neutral aspects of the same scene. Importantly, when neutral content is both central and peripheral, there is no difference in memory accuracy for the central and peripheral aspects of a scene (Kensinger, 2009). In parallel to evidence for directed attention leading to enhanced memory performance, when emotional features are present in a scene attention might be draw to those features, while other features are ignored, thereby increasing mnemonic accuracy for the emotional details relative to neutral backgrounds. For instance, when participants are asked to focus on the affect of a scene they show the typical memory trade-off, but when participants are asked to focus on the perceptual details of a scene the memory trade-off is diminished (Kensinger, Garoff-Eaton, Schacter, 2007). While arousal is certainly an important factor that influences attention and subsequent memory, it can only explain some, but not all, memory for the details of an experience (Kensinger, 2009). As described earlier, memory for details of our experiences is thought to be a crucial component of our subjective recollective experience and memory accuracy.

Valence and memory. Though the emotion-induced memory trade-off has mainly been explored with negative central elements and neutral peripheral elements, several other studies provide evidence that when arousal is controlled across positive and negative stimuli, memory for the details of an event is dependent upon the positive or negative valence of a stimulus (Ochsner, 2000; Kensinger, Garoff-Eaton, Schacter, 2007; Kensinger, O'Brien, et al. 2007; Waring & Kensinger, 2009). All of these studies suggest that negative items tend to be correctly remembered and/or vividly remembered more often than positive or neutral stimuli. In other words, people tend to remember seeing negative stimuli, like a mutilated hand or a tornado, with more accuracy and vivid detail than positive stimuli, like a cute puppy or a plate of appetizing food.

This differential effect of valence on memory has been proposed to result from different mechanisms underlying the encoding and retrieval of negative and positive stimuli. Negative items were characterized by focal memory enhancements as a result of more sensory processing of the item's distinctive details (Mickley and Kensinger, 2008). In addition, negative items were thought to recruit mechanisms that enhance distinctive encoding (Ochsner, 2000). For instance, negative items are known to have more general attention and perceptual biases, e.g. weapon's focus (Christianson & Failman, 1990, as cited in Ochsner, 2000). In addition, people tend to ruminate and cognitively elaborate more often about negative relative to positive information (Skowronski & Carlston, 1989; Thomas & Diener, 1990, as cited in Ochsner, 2000). Finally, negative stimuli are thought to contain more survival relevant information than positive stimuli (Ochsner, 2000).

Like negative items, positive items show a similar emotion-induced memory trade off relative to neutral items (Waring & Kensinger, 2008). Studies have found that positive items are more often subsequently recollected as just 'familiar' rather than vivid 'remembered' (Dewhurst & Conway, 1994; Ochsner, 2000). Neuroimaging studies have shown that positive stimuli later recollected as 'known' recruit regions involved in conceptual and self-referential processing (Mickley and Kensinger, 2008). This neuroimaging evidence suggest that some influences of valence on memory might be attributable to processing differences during encoding of positive and negative stimuli (Kensinger, 2009). Several studies have shown that experiencing positive affect increases the breadth of attentional selection (Rowe, Hirsh, & Anderson, 2006; Gable & Harmon-Jones, 2008; for review see Fredrickson, 2001). Although, to date no studies have directly tested how attentional resources are allocated for positive stimuli that are later accurately remembered and vividly recollected. Based on the behavioral and neural evidence for positive stimuli recollected as 'known', it could be the case that recollective experience for positive stimuli is influenced by different attentional allocation patterns than negative stimuli. Thus, the allocation of attentional resources for items correctly remembered with a vivid recollection of the episodic context may differ as a function of valence and not only arousal.

Pupillary arousal, eye-behaviors, and memory

Referring to the eyes, the 16th century french poet Guillaume du Bartas wrote “these lovely lamps, these windows of the soul” (1578). The notion that the eyes contain information not immediately available to our conscious mind has been a topic of poets, playwrights, and philosophers for centuries. Since the beginning of cognitive psychology, researchers have been exploiting this notion to gain access to the contents in the conscious and unconscious mind through the pupil response (Hess & Polt, 1960; Libby, Lacey, and Lacey, 1973 as cited in Bradley, Miccolo, Escrig, & Lang, 2008), and more recently through eye movements and fixations, as measures of attentional allocation (Sharot, Davidson, Carson, & Phelps, 2008; Smith & Squire, 2008; Hannula & Ranganath, 2009). While these projects have used various measures of eye-behavior, none of them have used both pupillometry and eye movements to assess the extent of information a person is acquiring and processing during encoding, as well as, the extent of cues subsequently used to retrieve the initial experience. By employing both measures of pupillometry and eye movements simultaneously, we hope to determine if physiological arousal during encoding is related patterns of attentional allocation that subsequently predict enhanced memory.

Measurement of emotional valence and arousal has been common ground for many debates as to the best indices to measure the nervous systems that drive emotional expression and enhancement of memory. Emotional ratings are by default highly subjective. Psychophysiological measures are objective gauges of emotional arousal (Lang et al., 1998). Drawbacks are inherent in most psychophysiological measurements of emotional arousal. Galvanic skin conductance is the most prominent psychophysiological index of emotional arousal, but suffers from several drawbacks, including a slow moving response and sensitivity to respiration (Andreassi, 2000).

Bradley and colleagues (2008) provided evidence that the pupil's response when observing affective pictures reflects emotional arousal and is associated with increased sympathetic activity. They found that the pupil response returns to baseline more quickly during exposure to positive and negative stimuli relative to neutral stimuli. In addition to the pupil response, they also measured galvanic skin conductance; a widely used, standard measure of sympathetic nervous system activity, and heart rate, a measure of parasympathetic nervous system activity. To determine whether the difference in pupil response could be attributed to sympathetic or parasympathetic activity, they correlated the activity between their different measures of physiological arousal. They found that the pupil response corresponds best with the skin conductance response. Thus, they attributed differences in the pupil response, particularly dilation of the pupil as it returns to baseline, to activation of the sympathetic nervous system. Importantly, Bradley and colleagues took many measures to ensure that other factors known to significantly influence pupil size are controlled, such as environmental luminance, stimulus luminance, and cognitive load.

In addition to measuring sympathetic nervous system activity, pupil responses have also been shown to differ as a function of whether a stimulus has been previously experienced. Similar to eye movements, increased pupil diameter relative to a neutral baseline has been shown to predict subsequent memory for novel images relative to old images in patients with unilateral MTL damage (including the amygdala; Laeng et al., 2007). This finding is evidence that pupillometry, like eye movements, may be another measure of recognition accuracy like subjective recognition responses. Given this evidence, it is important to consider complimentary findings that patients with unilateral or bilateral amygdala lesions. These patients do show impairments in emotional memory enhancement, but do not show impairments in subjective ratings of valence, arousal, or other measures of physiological arousal, like skin conductance (reviewed in Hamann, 2008). Thus, increased physiological arousal, measured by the pupil response, can still be attributed to modulation of mnemonic mechanisms at encoding through other sources in patients with amygdala damage. Further exploration of how pupil responses predict prior exposure to a stimulus, as well as the paricipant’s recollective experience, is needed to clarify which emotional arousal mechanisms relate to subsequent memory in normal participants.

Aims and Hypotheses

The current study intends to test the extent to which subsequent free recall, objective memory accuracy, and the subjective feeling-of-remembering are influenced by how attention is allocated when encoding stimuli that vary in the amount and type of salient features. In addition, we tested whether these forms of mnemonic experience manifest particular attention allocation patterns associated with recognition of previously experienced and new stimuli. The current study's objective was to replicate the findings of previous work, which has demonstrated more clustered attention is associated with the subjective feeling of remembering and stimuli high in arousal (Sharot et al., 2008). We applied the same experimental design as Sharot and colleagues (2008) and optimized the design to assess the attentional allocation for accurately recognized versus forgotten items and freely recalled versus not recalled items. Beyond replication, we also extended their design to address remaining questions concerning the attentional allocation for remembered items as a function of valence, through the including positive stimuli. We also tested whether physiological arousal, measured by the pupil response, during encoding and retrieval is associated subsequent memory.

Subsequent memory. In general, we expected to replicate the findings of many previous behavioral studies that show an emotional memory effect, in which memory is enhanced for emotionally arousing stimuli relative to neutral stimuli (Ochsner, 2000; Sharot, Delgado, Phelps, 2004). We expected to have preferential emotional memory enhancement for items that are subsequently recollected as 'remembered,' but no emotional memory enhancement for items subsequently recollected as 'familiar.' In addition, we expected to see similar emotional memory effects as a function of an item's subjective and physiological arousal.

Encoding. Based on Bradley and colleagues' (2008) study of the pupil's response to emotional pictures, we expected to find no difference between emotional and neutral items in pupil response during the initial light reflex (0-1 sec after stimulus onset) and more average pupil dilation for emotional items than neutral items during the 1-3 seconds after stimulus onset. The initial light reflex is the initial constriction of the pupil as it adjust to brief changes in light; usually occurring in the initial second after any brief change in light or stimulus (Bradley, Miccolo, Escrig, & Lang, 2008). Based on Sharot and colleagues' (2008) findings of an effect of emotional arousal on the clustering of eye-fixations and the Easterbrook hypothesis (1959), we anticipate that emotional items, regardless of valence, will show more clustering of fixations than neutral items. In the same vein, we expected to find that the span of attended cues (distance between fixations) will be more clustered or limited for items accurately remembered, vividly recollected as 'remember', and freely recalled.

For emotions effect on the sampling frequency, two opposing predictions can be derived from the Easterbrook's cue-utilization hypothesis and Sharot et al's findings. The cue-utilization hypothesis would predict that emotional items would be sampled less frequently than neutral items at encoding, although Sharot et al. found that emotional items were sampled more frequently than neutral items. Similar to Sharot et al's findings, we expected that the number of fixations per picture will differ as a function of subjective and physiological arousal, with more fixations for arousing items than neutral items. Finally, the pupil's response to items during encoding that are subsequently recalled, accurately remembered, vividly recollected, or forgotten has not been examined empirically. Based on the idea that encoding mechanisms are influenced by emotional arousal, we expected that pupil responses will return to baseline faster for later recalled, accurately remembered, and vividly recollected items.

Retrieval. We expected to see the same effect of stimulus arousal on the pupil response as at encoding with more average pupil dilation for emotional items than neutral items during the 1-3 seconds after stimulus onset. This finding would indicate that regardless of whether a stimulus had been previously experienced, it would still be emotionally arousing.

Attention allocation hypotheses at retrieval are based on the idea that attention is allocated in search of mnemonic cues in the stimulus to fulfill the recognition task, rather than as indicative of more or less cognitive elaboration on the features of a scene. Thus, we expected the span of attended cues will be more clustered or limited for items accurately remembered, vividly recollected as 'remember', and freely recalled due to less need to extensively search the image for mnemonic cues. This would be explained by stronger memory for part of the scene rather than the scene as a whole. Based on Sharot et al's findings, we predicted that the number of fixations per picture would differ as a function of subsequent memory accuracy, recollective experience, or free recall, with more fixations for forgotten, 'familiar', and non-recalled items. This finding would suggest that participants sample pictures with less strong memory representations more in their search for recollective cues before settling on a particular judgment.

Method

Participants

Twenty-seven healthy young adult volunteers (ages 19-30, *M*=21.4 ± 3.5) were recruited from the psychology department participant pool and posted advertisements. Seven participants were eliminated from analysis due to insufficient pupil and fixation measurements (<80% valid measurements). The remaining 20 participants (10 female) were included in the analysis. Eye measurements were measured for all 20 participants at encoding and retrieval. Participants were compensated with either class credit or $10 per hour for their time. The Emory University Institutional Review Board approved the study's procedures. All participants gave informed consent prior to participation in the experiments.

Stimuli

The stimulus set included 3 different sets of 90 pictures (270 pictures total) taken from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2008) and from our own set of emotional and neutral pictures. Pictures were equated on the presence of humans, visual complexity, and emotional arousal of positive and negative stimuli. Each set of 90 pictures consisted of 30 positive, 30 negative, and 30 neutral pictures (90 of each valence in total). Half of the images contained humans (135 without humans). All of the images were displayed in 16-bit grayscale and each image was modified such that mean luminance of the image matched the average luminance derived from all pictures in their default grayscale (Adobe Photoshop; version 7.0, Adobe Systems Inc., San Jose, CA). Bradley and colleagues (2001) showed that subjective and physiological arousal does not differ between color and grayscale photos. All selected pictures were in landscape orientation to ensure that they could be displayed on the full screen with no distortion. The pictures were 12 x 12 inches when displayed on the entire screen[1]. This size of image was deemed appropriate based on evidence demonstrating that physiological arousal responses are largest to pictures with similar dimensions to our stimuli (15” x 11”) relative to smaller images (5” x 3.5’; Codispoti & De Cesarei, 2007).

All stimuli were rated for valence, arousal, and visual complexity prior to the study in a separate group of similar participants (N=6). Valence was rated on a 1 (Negative) to 5 (Positive) scale with 3 categorized as neutral. Neutral pictures were rated as neutral (M=3.04, SEM=0.03). Emotional pictures were rated as positive (M=3.92, SEM=0.05) and negative (M=1.72, SEM=0.05). Arousal was rated on a scale from 1 (calm) to 5 (aroused). Neutral pictures were rated as low in arousal (M=1.88, SEM=0.05), while emotional pictures were rated as highly arousing (positive: M=2.93, SEM=0.04; negative: M= 3.25, SEM=0.07). Arousal ratings for positive and negative were both greater than for neutral pictures, t(5)= 6.7, p=0.001; t(5)=12.47, p=0.000, respectively. Arousal ratings were not different between positive and negative pictures, t(5)=2.03, p=0.10.

Participants in the current study rated the stimuli for valence and arousal similarly to the pilot study group. Valence and arousal were rated on the same scale as in the pilot group. Neutral pictures were rated as neutral (M=3.01, SEM=0.03). Emotional pictures were rated as positive (M=3.64, SEM=0.12) and negative (M=2.01, SEM=0.09). Neutral pictures were rated as low in arousal (M=1.93, SEM=0.12), while emotional pictures were rated as highly arousing (positive: M=2.85, SEM=0.13; negative: M= 3.21, SEM=0.13). Arousal ratings for positive and negative were both greater than neutral, t(19)= 9.6, p=0.000; t(19)=11.4, p=0.000, respectively. Negative stimuli were more arousing than positive stimuli, t(19)=6.6, p=0.000.

Behavioral Task

Before presenting each set of pictures at encoding or retrieval, participants situated themselves comfortably in a stationary chair, 60 cm from the screen. The eye tracker was calibrated to the orientation of the participant's eyes with an array of 9 dots placed across the extent of the screen. After calibration, fixation accuracy and consistency was tested by asking participants to look at a particular colored dot among a random scatter of dots that appeared across the screen.

Encoding. During the encoding session, participants were asked to passively watch 4 lists consisting of 45 pictures each. . Each list also contained 1 neutral buffer image at the beginning and end of the set to reduce primacy and recency effects (8 total). During encoding there were a total of 188 items; 60 were positive, 60 were negative, and 60 were neutral. This incidental encoding task allows participants to view the pictures naturally, rather than asking participants to distinguish characteristics of the picture, which might bias their eye movements. After an instruction screen, each picture was presented for 2 seconds (eye movements and pupillometry were measured continuously), after which a grayscale fixation screen appeared for 8 seconds. The fixation screens duration was 8 seconds to allow the pupil response to fully return to baseline before the presentation of a new stimulus. To keep the luminance of the stimuli consistent throughout runs and to reduce the initial light reflex, the fixation screen was modified to exactly match the luminance of all of the stimuli. If the participant successfully completed all 4 sets of pictures with at least 80% valid eye-movement measurements, they were asked to come back for a second session 1 week after the encoding session. Participants remained naïve to the subsequent memory tests in the second session (see Figure 1a).

Retrieval. One week after the initial encoding session participants were given a free recall and a remember/know recognition test. Participants first completed the free recall task in which they were asked to recall descriptions of the pictures they had seen the previous week. After participants recalled all of the pictures they could within 10 minutes, they were given the remember/know recognition test. In this type of recognition test participants are shown a mixture of old and new images in 6 sets of 45 images (180 old and 90 new; 270 total; 60 old items/valence, 30 new items/valence). Old and new stimulus sets were counterbalanced across participants.

For each picture presented the participant was asked to classify the pictures into vividly “remembered” that evoke a specific memory of an episodic context, which is a thought or feeling associated with the previous experience of the picture, simply “familiar” items that were “just known” (no episodic context) to have been experienced earlier, or new. Before the recognition test, participants were given instructions and examples for each response classification. They were then asked to repeat the criteria for each category (remember, familiar, and new) back to the experimenter in their own words to ensure they understood the instructions. After the experimenter checked the participant’s understanding, participants completed a practice session with the 8 neutral buffer images from the encoding session as stimuli.

After an instruction screen, each picture was presented for 2 seconds (eye movements and pupillometry were measured continuously), after which a grayscale fixation screen appeared for 2 more seconds, immediately followed by a self-paced response screen. A 1 second fixation screen followed the participant's response so that the participant could correct any initial response errors (see Figure 1b).

Ratings. Subjective ratings of valence and arousal were made for all stimuli during the retrieval session immediately following the remember/know recognition test. As described above, participants rated valence and arousal on a 5-point Likert-type scale, with valence ranging from 1 (negative) to 5 (positive) and arousal spanning from 1 (calm) to 5 (highly arousing). A valence rating of 3 denoted a neutral picture. Stimuli were presented and responses were recorded using the Psyscope X (B 53, Carnegie Mellon University). Stimuli were presented for 2 seconds, after which a self-paced valence rating response screen and the numbers 1 to 5 appeared. After the participant entered a valence response, the screen font color changed and another self-paced response screen asking for an arousal rating with the numbers 1 to 5 appeared.

Apparatus

Eye-tracking, pupillometry, and stimulus presentation was performed with the Tobii T120 eye-tracker system with an integrated 17" monitor (12.02 W x 12.02 H; 1280 x 1024 pixels; 94.25 pixels per inch; Tobii Technology). The eye-tracking system was located in an experimental room with no windows to control ambient light across conditions. A torchiere floor lamp with 110-watt light bulb was placed behind the participant to ensure consistent ambient light conditions during pupil response measurements.

Recording Eye Behavior Data

Pupil responses and eye movements were measured using both bright and dark pupil tracking. Dark pupil tracking detects the dark ellipse of the pupil within the iris, whereas bright pupil tracking detects the reflection of infrared light reflecting off the retina. Using an infrared light embedded around two cameras at the bottom of the eye-tracking monitor to create a corneal reflection, the eye-tracker captured the relative position and absolute size of the participant's pupil. As pictures appear on the entire screen (size: 12 x 12 in), eye-moments (in x, y coordinates of pixels) and pupil size (in mm) were measured at a rate of 120 Hz. Measurements were made continuously throughout each set of pictures at study and test. Participants were excluded from analyses if they had less than 80 percent valid measurements over a session. Participants were told prior to both the encoding and retrieval session that the eye-tracker was only being used to measure their pupil responses to control for any subjective bias in eye-gaze patterns.

Pilot study to calibrate experimental parameters. In order to determine the best design parameters for measuring effects of emotional arousal on pupil size, two pilot experiments were performed. The stimulus window and inter-trial interval were adjusted across the two experiments. The aim of the first pilot study (3 participants) was to replicate Bradley and colleagues' (2008) findings with the same design criteria (6-second stimulus presentation). The aim of the second pilot study was to determine whether a 2-second stimulus presentation window would elicit the same effect of emotional arousal on pupil size during and after the 2-second stimulus presentation. If the effect persisted after the 2-stimulus presentation design parameters could be the same as in Sharot et al. (2008) study, while still measuring the effects found in Bradley et al. (2008). Each design consisted of the same 96 pictures used in Bradley et al. with all pictures equated to the same luminance, which was matched to the mean luminance of all images in their default grayscale luminance. In all designs, the pupil size data from the period between 2-6 seconds was analyzed. In the second pilot study, the initial light reflex occurred for approximately 1 second. Results demonstrated an effect of emotional arousal on changes in pupil size during a 6- and 2-sec stimulus-presentation, replicating Bradley et al.'s results in both designs. Thus, a 2-second stimulus presentation was adopted for the current study; optimizing the design to replicate Sharot et al.'s design while still allowing adequate time to measure the pupil response.

Data Reduction

The Tobii fixation filter (Tobii Technology, Falls Church, VA) was used to transform the gaze positions into fixation points. The filter works under the assumption that fixations and jumps between fixations are the only eye motions present in the scan-path. First, the filter estimates the position of the fixations between changes in gaze position by using a median-estimator and then determines if adjacent fixations (in the time-domain) are closer together than a given spatial threshold. The algorithm used to detect fixations was as follows: if a segment of the gaze signal is of constant or slowly changing mean due to drift, it is classified as a fixation. Saccades were classified as abrupt changes in the gaze signal's mean. To determine if the mean is slowly changing or abruptly changing, the difference between the means of the sample points are compared within two sliding windows, which generates a difference parameter. This difference parameter is compared to the user-defined spatial threshold. This study used the default spatial threshold, which classifies any slow moving mean gaze point that is outside a radius of 35 pixels as a fixation. The estimate of the fixation's spatial coordinate (x, y in pixels) is calculated using the median of all samples in a candidate interval (interval in which the threshold is exceeded) and output in pixels.

After combining all of the data from each run in excel, the Pythagorean distance formula was used to calculate the distance in pixels between each fixation coordinate. Fixation coordinates were used in the distance formula if they occurred during the stimulus, thus excluding fixations that occurred when the stimulus changed or when the signal was lost. Then each inter-fixation distance value was summed and divided by the number of fixations during the 2-second window after stimulus onset to derive an average inter-fixation distance for each stimulus in the signal-processing program Acqknowledge (4.1, BIOPAC systems, Cambridge, England; see Figure 2).

For pupil measurements, samples where the pupil was obscured due to blinking or any other obstruction were identified by tobii studio as invalid samples. Using a MATLAB® script (2009a, The MathWorks, Natick, MA), invalid pupil size samples were linearly interpolated from the last valid measurement to the next valid measurement to approximate pupil size through out each time course. Using Acqknowledge, a 1-second pre-stimulus baseline was computed and subtracted from the measurements during the stimulus presentation. The minimal point of the average light reflex (Figure 4a) during encoding and retrieval occurred during the 1 second after stimulus onset. The pupil response to picture content was measured as the average pupil change (from baseline) during the 1 to 3 second period after stimulus onset. In this measurement window, the stimulus is visible for 1 second and the gray scale fixation screen is visible for an additional second (Fig. 1).

To ensure that all eye fixation effects were a not a function of the amount of time looking at each stimulus the average observation period for each participant was checked. All of the participants included in the analyses were looking at the image for 98% of the stimulus presentation period during encoding and 99% of the stimulus presentation period during retrieval.

Statistical Analysis

Statistically analysis was conducted using the average pupil change, the average number of fixations per condition during stimulus presentation (fixation frequency), and the average distance between sequential fixations over the stimulus presentation period (inter-fixation distance; for visual explanation see Fig. 2). Pupil change was implemented as a measure of sympathetic nervous system arousal. At encoding, the number of cues a participant fixates upon was indexed by fixation frequency. The span or range of cues the participant fixates upon was indexed by inter-fixation distance. A smaller inter-fixation distance would indicate more clustering of fixations for a particular condition. At retrieval, the number of cues needed to determine if a picture is “old”, “remembered”, “familiar”, or new was expressed by fixation frequency. The span or range of cues searched to determine if a picture is “old”, “remembered”, “familiar”, or new is expressed by inter-fixation distance. Thus, a smaller inter-fixation distance would indicate a smaller range of cues searched throughout the stimulus for a particular condition.

Repeated measures ANOVAs were used to test for main effects and interactions between emotion and memory. Effects sizes are reported as general eta squared (η_G^2), as it is the most appropriate statistic of effect size for repeated measures ANOVA (Bakeman, 2005). Similar to the general interpretation for η^2, interpretations of effect size for η_G^2 are as follows: effects of 0.02 to 0.13 are considered small, 0.13 to 0.26 are considered medium, and 0.26 and above are considered large. Paired t-test were used to test for differences between two within-subjects variables and post hoc pairwise comparisons of effects found in ANOVAs.

Results

Subsequent Memory

Effects of valence on memory. To test for an emotional memory effect, the mean proportions of corrected recognition and free recall were submitted to a 1 x 3 (positive, negative, and neutral valence) repeated measures ANOVA with planned comparisons. The corrected proportion of overall recognition was calculated by subtracting the proportion of hits minus the proportion of false alarms. There was an emotional memory effect for overall corrected recognition, $F(2,18) = 3.2$, $p=0.05$, $\eta_G^2 = 0.14$, with better corrected recognition for negative ($M=0.54$, $SEM=0.03$) than neutral items ($M=0.48$, $SEM=0.03$), $t(19)=2.39$, $p=0.03$, $d=0.54$ (Figure 3). The same calculation was used to derive the corrected recognition for items subsequently categorized as “remembered.” There was a emotional memory effect for corrected recognition of “remembered” items as well, $F(2,18) = 23.7$, $p=0.000$, $\eta_G^2 = 0.56$, with better corrected recognition for positive ($M=0.39$, $SEM=0.04$) than neutral ($M=0.32$, $SEM=0.04$), $t(19)=2.8$, $p=0.01$, $d=0.61$, negative (($M=0.50$, $SEM=0.03$) than neutral, $t(19)=6.9$, $p=0.000$, $d=1.5$, and negative than positive, $t(19)=3.9$, $p=0.001$, $d=0.89$ (Figure 3). To account for the non-independence of remember and familiar responses, the proportion of corrected recognition for "familiar" responses was calculated with the following formula: F=Fhit/(1- Rhit)-Kfa/(1-Rfa) (Yonelinas & Jacoby, 1994). There were no emotional memory effects for corrected recognition of items categorized as "familiar" (Figure 3). Thus, as expected, emotional items showed enhanced memory relative to neutral items. In particular, emotional items show enhanced vivid recollection relative to neutral items as evidenced by more 'remember' responses for emotional items.

The proportion of freely recalled items divided by valence was used to test for an emotional recall effect in a 1 x 3 (valence) repeated measures ANOVA. A significant emotional memory effect was observed, $F(2,18)=18.037$, $p<0.001$, $\eta_G^2 = 0.48$, with more items recalled for negative than positive or neutral items ($p<0.001$ corrected). However, it should be noted that the overall level of recall was very low, with some participants failing to recall any items from particular valence categories. Seventy percent of participants recalled at least 1 positive item ($M=1.4$, $SEM=0.33$), 95% participants recalled at least 1 negative item ($M=3.5$, $SEM=0.46$), and 55% participants recalled at least 1 neutral item ($M=1.2$, $SEM=0.32$). Table 1 depicts the mean and standard error of the mean (SEM) for overall recognition and recall (hits, false alarms, misses, correct rejections, and recalled) by valence.

Effects of arousal on memory. To test the hypothesis that memory was influenced by the subjective or physiological arousal of a stimulus, we performed a quartile split for each participant's arousal ratings and pupil responses during encoding. Because our stimuli were evenly split into 3 valence categories, two of which should be emotional arousing, we decided to split the subjective and physiological arousal by quartile rather than by median. We then compared the means of low, moderately low, moderately high, and high arousal items during encoding using repeated measures ANOVA with Bonferroni corrected post-hoc paired comparisons. The average median for subjective arousal across participants was 2, thus all items rated with an arousal rating of 3 and above were categorized as moderate high or high arousal and all items rated with an arousal rating of 2 and 1 were categorized as moderately low and low arousal. The average 1st quartile for physiological arousal was -0.42 mm, the average median for physiological arousal was -0.21 mm, and the average 3rd quartile for physiological arousal was 0.01 mm of average pupil change from baseline.

The dependent variable for recognition memory was the proportion of hits (items correctly categorized as old) during encoding. The ANOVA revealed a main effect of arousal quartile on proportion of hits, $F(3,17)=11.9$, $p<0.001$, $\eta_G^2 = 0.38$, with a higher proportion of hits for subjectively higher arousal items relative to subjectively lower arousal items. A linear contrasts reveals that the proportion of hits follows a linear trend, $F(1,19)=24.7$, $p=0.000$, $\eta_G^2 = 0.57$, with a higher proportion of hits for higher arousal items than lower arousal items (see Figure 5a). There was no difference in the proportion of hits between physiologically high arousal items and physiologically low arousal items. These results suggest that the memory accuracy was enhanced for stimuli perceived to be more emotionally arousing, but not necessarily for stimuli which are more physiologically arousing.

The dependent variable for recollective experience was the proportion of "remembered" or "familiar" hits. A 2 (recollective experience) x 4 (subjective arousal quartile) repeated measures ANOVA was used to test for an effect of subjective arousal on recollective experience accuracy. This analysis revealed an interaction subjective arousal and recollective experience, $F(3,17)=19.9$, $p<0.001$, $\eta_G^2 = 0.51$, with a higher proportion of "remembered" hits for subjectively higher arousal items relative to subjectively lower arousal items (see Figure 5b). The opposite trend was found for "familiar" hits, with a higher proportion of "familiar" hits for subjectively lower arousal items relative to subjectively higher arousal items. No differences were found in the proportion of "remembered" or "familiar" hits between physiologically high arousal items and physiologically low arousal items. These results suggest that the memory accuracy for vividly recollected items was enhanced for stimuli perceived to be more emotionally arousing, but not necessarily for stimuli which are more physiologically arousing. In contrast, memory accuracy for "known" or "familiar" items was worse for stimuli perceived to be more emotionally arousing, but not for stimuli which are more physiologically arousing.

The dependent variable for free recall was the proportion of items correctly recalled. There was a main effect of recall on proportion of items correctly recalled on subjective arousal, $F(2,18)=7.3$, $p<0.001$, $\eta_G^2 = 0.28$, with higher proportion of items correctly recalled for subjectively higher arousal items

(*M=0.06, SEM=0.01)* relative to subjectively lower arousal items (*M=0.02, SEM=0.01*). There was no difference in the proportion of items correctly recalled between physiologically high arousal items and physiologically low arousal items. These results suggest that recall was enhanced for stimuli perceived to be highly emotionally arousing, but not necessarily for stimuli which are more physiologically arousing.

Encoding

Pupillary Response. To test for an effect of emotional valence on the light reflex and pupil response, we performed a 1 x 3 (valence) repeated measures ANOVA. One-tailed paired *t*-tests were used to test for an effect of the directional hypothesis that the pupil response for emotional stimuli would return to baseline faster than the pupil response for neutral stimuli. We predicted no difference between emotional and neutral items in pupil response during the initial light reflex (0-1 sec after onset) and more average pupil dilation for emotional items than neutral items during the 1-3 seconds after stimulus onset. Our findings confirm our predictions. There was no difference in the pupil response between emotional and neutral items during the initial light reflex period (0-1 sec). There was an effect of valence on pupil response during the 1 to 3 window after stimulus onset, $F(2,18)=4.7$, $p=0.025$, $\eta_G^2 = 0.2$, with larger pupil dilation for positive than neutral, one-tailed $t(19)=1.9$, $p=0.04$, negative than neutral, one-tailed $t(19)=2.5$, $p=0.01$, and negative than positive, one-tailed $t(19)= 1.7$, $p=0.05$. These results demonstrate that emotional items (positive and negative) arouse the sympathetic nervous system more than neutral items during encoding. Figure 4a depicts the average pupil response over the 0 to 3 seconds after stimulus onset split by valence.

Eye movements. To test for an effect of emotional valence on eye movements (fixation distance and fixation frequency), we performed a 1 x 3 (valence) repeated measures ANOVA. We predicted that emotional items would show more clustering of fixations than neutral items and the average of fixations would be reduced more for emotional items than neutral items. This ANOVA revealed that inter-fixation distance was characterized by no effect of valence, *F*(2,18)=1.0, p=0.37, $\eta_G^2 = 0.05$, thus there was no difference in the average fixation distance for positive, negative, or neutral items. There was an effect of valence on fixation frequency, *F*(2,18)=3.7, p=0.03, $\eta_G^2 = 0.16$, with more fixations for positive (*M*=5.6, *SEM*=0.15) than neutral (*M*=5.4, *SEM*=0.19) items, *t*(19)=2.9, p=0.008. These results suggest that emotional scenes elicit enhanced sampling rates, but not enhanced clustering of eye fixations, relative to neutral scenes.

Effects of valence on pupillary correlates of subsequent memory. To determine if arousal-related pupil responses were associated with subsequent memory during encoding, we performed a 2 (subsequent memory) x 3 (valence) repeated measures ANOVA with Bonferroni corrected paired comparisons. We predicted that pupil responses would return to baseline faster for emotional items, as shown above, and items accurately remembered, vividly recollected as 'remember', and freely recalled.

For memory accuracy, there was a main effect of valence for pupil response, *F*(2,18)=3.73, p=0.046, $\eta_G^2 = 0.07$, with the pupil response returning to baseline faster for negative items than positive or neutral items. There was also a main effect of subsequent memory, *F*(1,19)= 4.7, p=0.044, η_G^2 =0.04, with the pupil dilating more for subsequently forgotten items than subsequently remembered items. As demonstrated above, emotional items elicit more arousal than neutral items. In addition, these results suggest that subsequently remembered items elicit less arousal at encoding than subsequently forgotten items.

For recollective experience, there was an interaction of valence and recollective experience, *F*(2,18)=4.33, p=0.02, η_G^2 =0.06, with larger pupil dilation for negative items vividly recollected as ‘remember’ than negative items recollected as ‘familiar,’ *t*(19)=2.5, p=0.02. This effect was not found for positive or neutral ‘remember’ or ‘familiar’ items. These results suggest that negative items which are subsequently vividly recollected might be more arousing than negative items which are subsequently familiar.

For freely recalled items there was main effect of recall, *F*(1,13)=6.9, p=0.02, η_G^2 =0.09, with a greater pupil response for recalled than not recalled items. When neutral items are also included in the ANOVA, there is a only trend toward a main effect of recall, *F*(1,7)=5.03, p=0.06, η_G^2 =0.08, with a greater pupil response for recalled than not recalled items.

Effects of valence on eye movement correlates of subsequent memory. To determine if the distance between fixations (inter-fixation distance) during encoding differed as a function of subsequent memory and picture valence, we performed a 2 (subsequent memory) x 3 (valence) repeated measures ANOVA with Bonferroni corrected paired comparisons. We predicted that the span of attended cues would be more clustered or limited for items that would subsequently be accurately remembered, particularly those items that would later be vividly recollected and be endorsed by participants with a ‘remember’ response. In addition, we predicted that attention would be more clustered for items that were subsequently recalled than not recalled.

For recognition accuracy, inter-fixation distance was characterized by a trend towards a main effect of valence, $F(2,18)= 3.07$, $p = 0.06$, $\eta_G^2 =0.04$, with numerically more clustered fixations for negative (M=223.2, SEM=5.2) and neutral (M=224.2, SEM=5.8) items than neutral (M=231.7, SEM=6.2) items. There was a main effect of valence when a 2 [emotional (positive and negative) vs. neutral] x 2 (hit vs. miss) repeated measures ANOVA was performed, $F(1,19)=6.53$, $p=0.02$, $\eta_G^2 =0.07$, with significantly more clustering of fixations for emotional than for neutral items.

For recollective experience, one outlying participant in the neutral ‘familiar’ hit condition (n=19) was removed from the following analyses. Inter-fixation distance was characterized by a significant interaction of valence and recollective experience, $F(2,17)=3.27$, $p=0.05$, $\eta_G^2 = 0.05$ (Figure 6a). To examine the components of this interaction, t-test were run for comparisons between inter-fixation distance for each valence as a function of whether the items were ‘remembered’ or ‘familiar.’ Inter-fixation distances were significantly more clustered in negative remembered items than neutral remembered items, $t(19)=2.76$, $p=0.01$. Inter-fixation distances were also significantly more clustered in negative remembered items than positive remembered items, $t(19)=2.09$, $p=0.05$. Inter-fixation distances were not statistically different between positive remembered items and neutral remembered items. In addition, inter-fixation distances did not differ as a function of valence for items subsequently recollected as familiar.

We also performed a 2 (negative only vs. neutral) x 2 (‘remember’ hit vs. ‘familiar’ hit) to replicate Sharot et al’s analysis. Inter-fixation distance was again characterized by a significant interaction of valence and recollective experience, $F(1,18)=4.25$, $p=0.05$, $\eta_G^2 = 0.07$, with more clustering for neutral ‘familiar’ hits than ‘remember’ hits and more clustering for negative ‘remember’ hits than ‘familiar’ hits. No main effects or interactions were found for ‘remember’ or ‘familiar’ items in a 2 (positive only vs. neutral) x 2 (recollective experience measures ANOVA.

To follow-up on the 3 (valence) x 2 (recollective experience) interaction, we also performed a 2 (recollective experience) x 2 (positive vs. negative) repeated measures ANOVA. When testing between the two emotional valences, inter-fixation distance was characterized by a significant interaction of valence and recollective experience, $F(1,19)=7.6$, $p=0.01$, $\eta_G^2 = 0.09$, with smaller clusters for positive ‘familiar’ hits than ‘remember’ hits and smaller clusters for negative ‘remember’ hits than ‘familiar’ hits.

Finally, to explore these interactions further with consideration of valence and high arousal, another 3 (valence; high arousal only for positive and negative) x 2 (recollective experience) repeated measures ANOVA was conducted. High arousal was operationally defined by a median split of each individual’s pupillary response during the 1-3 second window, using only those items which had a pupillary response above an individual’s median pupillary response. This ANOVA revealed the same interaction of valence and recollective experience (figure 6a), but with only high arousal positive and negative items instead of positive and negative items regardless of valence, $F(2,18)=4.26$, $p=0.01$, $\eta_G^2=0.06$. This ANOVA also showed a main effect of valence for highly physiologically arousing items according to the pupillary response, $F(2,18)=3.66$, with more clustered fixations for positive than neutral items ($p=0.03$ corrected; see figure 6b). When the same ANOVA was used to test minimally arousing items by valence (below the individual's median pupillary response), there was no effect of valence or interaction between valence and recollective experience. In summary, these results suggest that greater clustering of fixations during encoding was related to subsequent vivid recollection for negative items, whereas less clustering of fixations at encoding was related to subsequent vivid recollection of positive and neutral items. For items subsequently recollected without an episodic context, i.e. familiar items, there was no difference in the inter-fixation distance. In addition, more clustering of fixations at encoding was related to vivid recollection of highly physiologically arousing, emotional (positive and negative) items relative to vividly remembered neutral items.

For freely recalled items, inter-fixation distance was characterized by a main effect of recall, $F(1,7)=14.5$, $p=0.007$, $\eta_G^2=0.16$, with greater clustering of fixations for recalled than not recalled items. This finding tentatively suggests that attention was narrower for subsequently recalled items than items not recalled.

To determine if the number of fixations per picture (fixation frequency) during encoding differed as a function of subsequent memory and item valence, we performed a 2 (subsequent memory) x 3 (valence) repeated measures ANOVA with Bonferroni corrected paired comparisons. Based on Sharot et al's findings, we predicted that the number of fixations per picture would differ as a function of valence, with more fixations for emotional items than neutral items, but not necessarily as a function of subsequent memory accuracy, recollective experience, or free recall.

For memory accuracy, fixation frequency was characterized by a main effect of valence, $F(1,19)=4.4$, $p=0.02$, , $\eta_G^2 =0.09$, with more fixations for emotional than neutral. Post hoc comparisons show that this effect was driven by positive images having more fixations than neutral ($p=0.03$ corrected). There was also a reliable main effect of recognition accuracy, $F(1,19)=23.9$, $p=0.000$, $\eta_G^2 =0.21$, with more fixations per picture for subsequently remembered than forgotten items. These results compliment findings of the 1 x 3 (valence) ANOVA before which suggested that positive items are sampled more than neutral items at encoding. In addition, pictures which were accurately remembered are sampled more during encoding than items which were forgotten.

For recollective experience, fixation frequency was characterized by a main effect of recollective experience, $F(2,18)=4.5$, $p=0.047$, $\eta_G^2 = 0.06$, with more fixations for remembered than familiar items. There was no main effect of valence as was found in Sharot et al. (2008). We also performed a 2 (negative only vs. neutral) x 2 (‘remember’ hit vs. ‘familiar’ hit) to replicate Sharot et al’s analysis. In this analysis, fixation frequency was characterized by a trend toward a main effect of recollective experience, $F(1,19)=3.77$, $p=0.07$, $\eta_G^2 = 0.06$, with more fixations for remembered than familiar items. When the same 2 x 2 ANOVA was done with positive items only instead, fixation frequency was characterized by a main effect of recollective experience, $F(1,19) = 4.45$, $p=0.05$, $\eta_G^2 = 0.09$, with more fixations for remembered than familiar items. These results suggest that items which are subsequently vividly recollected are sampled more often than items recollected without an episodic context.

For freely recalled items, there were no main effects or interactions based on fixation frequency.

Relationship between arousal on eye movement correlates of subsequent memory. To determine if the distance between fixations (inter-fixation distance) during encoding differed as a function of subsequent memory and physiological arousal (measured with pupillometry), we performed a 2 (subsequent memory) x 4 (pupillary response quartile) repeated measures ANOVA with Bonferroni corrected paired comparisons. We predicted the span of attended cues would be more clustered or limited for more arousing items, as well as, items that were accurately remembered, vividly recollected as 'remember', and freely recalled. For recognition accuracy, there were no main effects or interactions based on inter-fixation distance. For recollective experience, inter-fixation distance was characterized by a main effect of arousal, $F(3,17)=3.9$, $p=0.01$, $\eta_G^2 = 0.06$, with more clustered fixations for higher arousal items (M=218 pixels, SEM= 7.9) than lower arousal items (M=232 pixels, SEM=7.2; p=0.02 corrected). Inconsistent with findings for valence and recollective experience, this result suggests that attention was only narrower for items that are more physiologically arousing, but not differentially narrower for 'remembered' versus 'familiar' items (see Fig. 9). For freely recalled items, there were no main effects or interactions based on inter-fixation distance, thus there was no difference in inter-fixation distance as a function of subsequent recall and physiological arousal.

To determine if the number of fixations (fixation frequency) per picture during encoding differed as a function of subsequent memory and physiological arousal, we performed a 2 (subsequent memory) x 4 (pupillary response quartile) repeated measures ANOVA with Bonferroni corrected paired comparisons. We predicted that the number of fixations per picture would differ as a function of arousal, with more fixations for emotional items than neutral items, but not necessarily as a function of subsequent memory accuracy, recollective experience, or free recall. For memory accuracy, fixation frequency was characterized by a main effect of arousal, $F(3,17)=3.5$, $p=0.02$, , $\eta_G^2 =0.07$, with more fixations for low arousal items than high arousal items. There was also a reliable main effect of recognition accuracy, $F(1,19)=22.4$, $p<0.001$, $\eta_G^2 =0.20$, with more fixations per picture for subsequently remembered ($M=5.6$, $SEM=0.15$) than forgotten items ($M=5.2$, $SEM=0.19$; $p<0.001$ corrected). For recollective experience, fixation frequency was characterized by a main effect of arousal, $F(3,17)=3.5$, $p=0.02$, $\eta_G^2 =0.04$, with more fixations per picture for low arousal pictures than high arousal pictures. For freely recalled items, there were no main effects or interactions based on fixation frequency. These findings suggest that sampling rates are reduced for highly physiologically arousing items relative to items that are not physiologically arousing.

To determine if the distance between fixations (inter-fixation distance) during encoding differed as a function of subsequent memory and subjective arousal, we performed a 2 (subsequent memory) x 4 (subjective arousal quartile) repeated measures ANOVA with Bonferroni corrected paired comparisons.

Similar to our prediction for valence, we predicted the span of attended cues would be more clustered or limited for more arousing items, as well as, items that were accurately remembered, vividly recollected as ‘remember’, and freely recalled. There were no main effects or interactions as a function of any of the memory or arousal categories.

To determine if the number of fixations (fixation frequency) per picture during encoding differed as a function of subsequent memory and physiological arousal, we performed a 2 (subsequent memory) x 4 (pupillary response quartile) repeated measures ANOVA with Bonferroni corrected paired comparisons. We predicted that the number of fixations per picture would differ as a function of arousal, with more fixations for emotional items than neutral items, but not necessarily as a function of subsequent memory accuracy, recollective experience, or free recall. For memory accuracy, fixation frequency was characterized by a main effect of recognition accuracy, $F(1,17)=27.5$, $p<0.001$, $\eta_G^2=0.22$, with more fixations per picture for subsequently remembered ($M=5.6$, $SEM=0.17$) than forgotten items ($M=5.2$, $SEM=0.19$; $p<0.001$ corrected). For recollective experience, fixation frequency was characterized by a main effect of arousal, $F(3,17)=3.5$, $p=0.02$, , $\eta_G^2=0.03$, with more fixations per picture for low arousal pictures than high arousal pictures. For recollected and freely recalled items, there were no main effects or interactions based on fixation frequency. In parallel to the results for fixation frequency as a function of valence these results suggest that accurately remembered items are sampled more at encoding than forgotten items.

Retrieval

Pupil Response. To test for an effect of emotional valence on the light reflex and pupillary response during retrieval, we performed a 1 x 3 (valence) repeated measures ANOVA. One-tailed paired *t*-tests were used to test for an effect of the directional hypothesis that the pupillary response for emotional stimuli would return to baseline faster than the pupillary response for neutral stimuli. We predicted no difference between emotional and neutral items in pupillary response during the initial light reflex (0-1 sec after onset) and more average pupil dilation for emotional items than neutral items during the 1-3 seconds after stimulus onset. There was a main effect of valence in the pupillary response during the initial light reflex period (0-1 sec), $F(2,18)=7.2$, $p=0.002$, $\eta_G^2 = 0.27$, with larger pupil dilation for positive than neutral, one-tailed $t(19)=3.4$, $p=0.004$ corrected, and negative than neutral, one-tailed $t(19)=3.4$, $p=0.004$ corrected. There was an effect of valence on pupillary response during the 1 to 3 second window after stimulus onset, $F(2,18)=20$, $p=0.000$, $\eta_G^2 = 0.51$, with larger pupil dilation for positive than neutral, one-tailed $t(19)=4.7$, $p=0.000$, and negative than neutral, one-tailed $t(19)=5.8$, $p=0.000$. Figure 4b depicts the average pupillary response over the 1 to 3 seconds after stimulus onset split by valence. These results demonstrate that emotional items (positive and negative) arouse the sympathetic nervous system more than neutral items during retrieval. In addition, there might be an effect of previous exposure on the pupillary response's light reflex latency, with the pupillary response beginning to return to baseline faster than neutral and before the typical trough of the light reflex. The trough of the light reflex should not differ between emotional categories because it should only be dependent upon the change in ambient light, which was controlled between stimuli.

Eye movements. To test for an effect of emotional valence on eye movements (fixation distance and fixation frequency) during retrieval, we performed a 1 x 3 (valence) repeated measures ANOVA. We predicted that emotional items would show more clustering of fixations than neutral items and that the number of cues attended to would be reduced more for emotional items than neutral items. Inter-fixation distance was characterized with no effect of valence, $F(2,18)=0.97$, $p=0.37$, $\eta_G^2 = 0.05$, thus there was no difference in the average fixation distance for emotional items and neutral items. There was an effect of valence on fixation frequency, $F(2,18)=5.7$, $p=0.007$, $\eta_G^2 = 0.23$, with more fixations for positive ($M=5.76$, $SEM=0.13$) than neutral ($M=5.61$, $SEM=0.15$) items, $t(19)=2.9$, $p=0.03$ corrected, and positive than negative ($M=5.4$, $SEM=0.19$), $t(19)=3.7$, $p=0.005$ corrected. These results parallel the findings during encoding which suggest that positive images are sampled more than negative or neutral images in general, regardless of whether the item was being encoded or retrieved.

Relationship between valence and pupillary correlates of subsequent memory. To determine if arousal related pupillary responses influenced subsequent memory during retrieval, we performed a 4 (subsequent memory) or 2 (recollective experience or recall) x 3 (valence) repeated measures ANOVA with Bonferroni corrected paired comparisons. We predicted that pupillary responses would return to baseline faster for emotional items, as shown above, and items accurately remembered, vividly recollected as ‘remember’, and freely recalled.

For memory accuracy at retrieval, as was found above, the pupillary response was characterized by a main effect of valence, $F(2,17)=9.6$, $p<0.001$, $\eta_G^2 = 0.08$, with the pupillary response returning to baseline faster for negative and positive items than neutral items. There was also a main effect of subsequent memory, $F(3, 16)= 4.1$, $p=0.03$, $\eta_G^2= 0.08$, with the pupil dilating more for items the participant subjectively thought were old (hits and false alarms) than items the participant subjectively thought were new (misses and correct rejections). This result suggests that items subjectively responded to as old elicited more physiological arousal than items subjectively responded to as new.

For recollective experience, consistent with the previous analyses, pupillary response was characterized by a main effect of valence, $F(2,18)=8.2$, $p=0.003$, $\eta_G^2 = 0.15$, with larger average pupil dilation for positive and negative items than neutral items. Pupil response was also characterized by a main effect of recollective experience, $F(1,19)=19.33$, $p<0.001$, $\eta_G^2 =0.14$, with larger average pupil dilation for vividly recollected items than items subsequently recollected as just ‘familiar’. This finding suggests that vividly recollected items elicit more physiological arousal than items recollected without an episodic context.

For freely recalled items pupillary response was characterized by a main effect of recall, $F(1,7)=9.2$, $p=0.02$,$\eta_G^2 =0.18$, with a greater pupillary response for recalled than not recalled items. As described above, this finding was tentative due to the low degrees of freedom, but suggests that the arousal response of the pupil might be a marker for strong memory of a previously experienced stimulus.

Effects of valence on eye movement correlates of subsequent memory. To determine if the distance between fixations (inter-fixation distance) during retrieval differed as a function of subsequent memory and picture valence, we performed either a 4 (subsequent memory) or 2 (recollective experience or recall) x 3 (valence) repeated measures ANOVA with Bonferroni corrected paired comparisons. We predicted the span of attended cues would be more clustered or limited for items accurately remembered, vividly recollected as 'remember', and freely recalled due to less need to extensively search the image for memories cues.

For recognition accuracy (4 x 3 ANOVA), no main effects or interactions were found for inter-fixation distance. For recollective experience, inter-fixation distance was characterized by an interaction of valence and recollective experience, $F(2,18)=3.3$, $p=0.05$, $\eta_G^2 = 0.09$, with smaller clusters for negative remembered hits than familiar hits $t(19)=4.6$, $p<0.005$ (Figure 7a). This ANOVA also revealed a main effect of recollective experience, $F(1,19)=16.03$, $p=0.001$, $\eta_G^2 = 0.13$, with smaller clusters for remembered hits than familiar hits. We also performed a 2 (negative only vs. neutral) x 2 ('remember' hit vs. 'familiar' hit) to replicate Sharot et al's analysis. This replication analysis yielded an interaction of valence and recollective experience, $F(1,19)=5.1$, $p=0.04$, $\eta_G^2 = 0.20$, with smaller clusters for negative 'remember' hits than neutral, $t(19)=2.9$, $p=0.009$. Within the each valence there were only smaller clusters for 'remember' hits than 'familiar' hits for negative, $t(19)=4.64$, $p<0.001$. No main effects were found for 'remember' or 'familiar' items in a positive only versus neutral analysis.

There were no main effects or interactions for inter-fixation distance as a function of recall and valence. These results suggest that enhanced clustering of fixations during retrieval was indicative of vivid recollection of only negative items relative to negative items recollected with no episodic context. While enhanced clustering of fixations during retrieval of positive and neutral images does not differ as a function of recollective experience.

To determine if the number of fixations per picture (fixation frequency) during encoding differed as a function of subsequent memory and item valence, we performed a 4 (subsequent memory) or 2 (recollective experience or recall) x 3 (valence) repeated measures ANOVA with Bonferroni corrected paired comparisons. Based on Sharot et al's findings, we predicted that the number of fixations per picture would differ as a function of subsequent memory accuracy, recollective experience, or free recall, with more fixations for forgotten, 'familiar', and non-recalled items, but not necessarily as a function valence. This finding would provide evidence of more sampling of the image in a search for memory cues.

For memory accuracy, fixation frequency was characterized by a main effect of valence, $F(2,18)=3.81$, $p=0.03$, $\eta_G^2 = 0.04$, with more fixations for positive than negative or neutral items. Post hoc comparisons show that this effect was driven by positive images having more fixations than negative ($p=0.01$) and neutral ($p=0.03$). This effect mirrors findings from previous analysis and suggests again that positive items are sampled more often than negative or neutral items at encoding and retrieval.

For recollective experience, fixation frequency was characterized by an interaction of valence and recollective experience, $F(1,19)=3.5$, $p=0.04$, $\eta_G^2 =$

0.07, with more fixations for negative familiar than remembered items. We also performed a 2 (negative only vs. neutral) x 2 (‘remember’ hit vs. ‘familiar’ hit) to replicate Sharot et al’s analysis. In this analysis, fixation frequency was characterized by an interaction of valence and recollective experience, $F(1,19)=11.86$, $p=0.003$, $\eta_G^2 = 0.10$, with more fixations for negative familiar than negative remember items. This finding suggests that the sampling rate was enhanced when searching for recollective cues for negative items which lack an episodic context and was diminished when searching for recollective cues for vividly recollected items. When the same 2 x 2 ANOVA was done with positive items instead, fixation frequency was characterized by a main effect of valence, $F(1,19) = 6.02$, $p=0.02$, $\eta_G^2 = 0.09$, with more fixations for positive than neutral items. This finding suggests that the sampling rate was enhanced when searching for any positive item relative to neutral items regardless of how it was subsequently recollected.

For freely recalled items, there were no main effects or interactions based on fixation frequency, thus there was no difference in fixation frequency as a function of subsequent recall and valence.

Discussion

The results of this study partially supported the hypothesis that the distribution of attention during encoding would be narrower for accurate, vivid recollections of stimuli, regardless of valence. The results also partially support the cue utilization hypothesis (Easterbrook, 1959), which posits that attention narrows as a function of emotional arousal. We found that memory performance, recall, and vivid recollection were related to more narrowing of attention for negative pictures than positive or neutral pictures during encoding. During retrieval, vivid recollection may be prompted by enhanced memory for a few salient features of the pictures, regardless of valence, but with more fixations to positive pictures. Finally, the current findings supported the hypothesis that the pupillary response would increase with the emotional arousal of a stimulus and that higher physiological arousal during encoding and retrieval would be associated with subjectively vivid memories.

Pupillary response indexes emotional arousal and subsequent vivid recognition memory and recall

As predicted, pupillary changes were larger after the initial light reflex when viewing emotionally arousing pictures during encoding and retrieval (Bradley, Miccolo, Escrig, & Lang, 2008). We found that negative items were more physiologically arousing than positive items and positive items were more arousing than neutral items after the initial light reflex. This pattern of findings mirrors the average subjective arousal ratings as well as the overall pattern of memory enhancement, especially for later 'remembered' stimuli. This finding gives support to previous studies that find increased pupil dilation to emotional stimuli (Hess & Polt, 1960; Libby, Lacey, and Lacey, 1973 as cited in Bradley et al., 2008). To our knowledge, this study is the first to document that pupil size increases as a function of emotional arousal and valence during retrieval of emotional stimuli.

In addition to indexing emotional arousal, larger pupillary responses during encoding were associated with negative, vividly recollected items and subsequently recalled items. During recognition, larger pupillary responses were associated with vividly recollected items and items subjectively recognized as old (e.g. hits and false alarms). Thus, physiological arousal measured by the pupillary response is associated with vivid memory at encoding and retrieval. Other studies have documented similar, but slightly different results. Laeng et al. (2006) found that increased pupil diameter relative to a neutral baseline during retrieval was associated with viewing novel images relative to previously seen images in patients with unilateral MTL damage (including the amygdala) and no explicit awareness of 'knowing' the pictures already. To our knowledge, the current study is the first to examine whether the pupillary response differs as a function of memory in normal healthy volunteers, but also the first to find that larger pupillary responses during encoding and retrieval are associated with vivid, but not necessarily accurate, recollection.

Enhanced memory as a function of higher subjective arousal

As expected the current study replicated the typical observed experimental emotional memory enhancement for overall recognition accuracy and recognition remember judgments. Numerous studies have found that participants show enhanced memory for emotional relative to neutral items (for review see LaBar and Cabeza, 2006). Several other studies demonstrate that participants show enhanced vivid recollection for emotional relative to neutral items as evidenced by more accurate 'remember' responses than 'know' for emotional items (Oschner, 2000; for review see Heuer & Reisberg, 1992). Here we found that memory accuracy and the recollective component of recognition memory were both enhanced for stimuli subjectively rated as highly emotionally arousing. However, when emotional arousal was defined on the basis of pupillary arousal responses, only recall was predicted by larger pupillary responses. Several studies have reported that subjective and physiological measurements of arousal tend to correlate only moderately, so that it is not unexpected that the subjective arousal ratings and pupillary arousal measures were only moderately correlated with each other (r=0.11; e.g., Lang, Greenwald, Bradley, Hamm, 1993). In addition, subjective arousal might be taking into account other cognitive factors like how interesting the pictures is to the participant, thus explaining more variance in the memory performance than just physiological arousal. Future studies of emotional memory should utilize pupillary measures of arousal concurrently with other measures of physiological arousal, like galvanic skin conductance, to determine if pupillometry is a sensitive enough measure of arousal to consistently relate to enhanced memory performance for emotionally arousing stimuli.

Distribution of attention during encoding predicts memory as a function of valence

Attention was differentially allocated according to stimulus valence during encoding. Positive stimuli were sampled more frequently than negative or neutral stimuli during encoding. Eye fixations were more clustered during encoding of negative remembered pictures relative to negative familiar pictures, replicating previous results from Sharot et al. (2008). In contrast, eye fixations were not differentially clustered during encoding for positive and neutral remembered pictures relative to positive and neutral familiar pictures. These findings suggest that vivid recollection is related to encoding of a few distinct features of highly arousing negative pictures, but for positive, low arousal negative, and neutral stimuli there is no difference between attention allocation for the ‘remember’ and 'familiar' component of recollective experience. Thus, we did not fully replicate the findings of Sharot and colleagues (2008), which suggested that more clustering of eye fixations during encoding was independently indicative of emotional arousal and later vivid recollection. Rather we find that narrower attention during encoding is dependent upon emotional valence, emotional arousal, and subsequent recollective experience. Importantly, the pattern during encoding for negative later vividly recollected items was the same across both this study and Sharot et al. (2008), with more clustering of fixations for highly arousing negative 'remembered' stimuli than highly arousing negative 'familiar' stimuli. The main comparable difference is in the allocation of attentional resources for neutral stimuli, which showed a non-differential pattern for recollective experience in our study relative to theirs, with no difference in clustering of fixations for neutral 'familiar' stimuli than 'remembered' stimuli. In addition, our inclusion of the pupillary response as a measure of physiological arousal and subjective ratings of arousal allowed us to investigate the contribution of physiological arousal to this narrowing of attention effect. The current study only found narrower attention for highly arousing negative items that were later vividly recollected. Interestingly, the same pattern of attention allocation as neutral stimuli was found in our study for positive stimuli, which was not examined in Sharot et al. (2008).

The inconsistencies of between this study and Sharot et al. (2008) could be attributable to minute differences in stimulus features, like color versus grey scale or stimulus size. The lack of color in our stimuli might have reduced the distinctiveness of each of the images making it more difficult to dissociate salient and non-salient features than in Sharot et al's colored stimuli. Though this is a possibility, the pictures were at least matched for visual complexity across valence categories, like Sharot's stimuli, which is related to how many distinctive features a picture could have. Stimulus size might have increased the arousal of the stimuli relative to Sharot's study (Codispoti & De Cesarei, 2007). But it is unlikely that stimulus size significantly affected our results because the arousal of the stimuli was matched accordingly for emotional and neutral stimuli, just as in the Sharot et al. study. Differences in stimulus size also increased the area that needed to be covered to attend to separate aspects of the stimulus. This might have affected the amount of information that could be gathered from fixating on one point of the screen. Differences in size are also unlikely to explain our differing results because the magnitude of our average fixation distances were scaled up according to the same 2:1 size ratio between our stimuli and their stimuli. In addition, stimulus size did not affect the average number of fixation per stimulus, as both studies have similar fixation frequencies. Finally, the duration of stimulus presentation was designed to exactly replicate Sharot et al's design. The only major difference in design was the delay between encoding and retrieval. Sharot and colleagues had only a 45-minute delay, while our study used a 7-day delay. Testing delay does not affect our results due to the observation that memory accuracy levels were similar for our study and their study. Thus, because differences in stimulus features or presentation do not account for the differences in results between our study and Sharot's et al., our study is probably demonstrating theoretically relevant results.

The finding that the distribution of attention for items later vividly recollected is dependent upon valence supports and extends the cue-utilization hypothesis proposed by Easterbrook (1959). This finding might be explained by differences in the distribution of distinctive and salient features intrinsic to the valence of the stimulus. For instance, negative images might naturally tend to have salient information concentrated in one part of the stimulus, while positive and neutral images might naturally tend to have salient features more dispersed throughout the image. This interpretation is speculative and has not been empirically examined, although we did use subjective ratings of visual complexity to control for visual complexity across all valences. Future studies could assess the distribution of content in regularly used emotional pictures.

Another similar explanation, more supported by previous research, is that the strength by which attention is captured by salient features of a stimulus during encoding is dependent on the stimulus valence and arousal. For instance, negative images are known to be associated with several strong attentional and perceptual biases, like weapons focus (Loftus, Loftus, & Messo, 1987; Kramer, Buckhout, Eugenio, 1990; for a meta-analysis of weapons focus see Steblay, 1992; Ochsner, 2000). In addition, negative images are thought to contain more survival relevant information (Ochsner, 2000; Hamann, 2001). Also, less research has found similar perceptual or attentional biases to positive and neutral images, although memory researchers have proposed that remembering positive information requires less detailed, conceptual processing and found that positive information is more often ‘familiar’ than ‘remembered’ (for review see Kensinger, 2009). The current study demonstrated that positive and neutral stimuli later recollected as ‘familiar’ are not associated with focusing on a particular distribution of the features in the stimulus during encoding. We speculate that the best explanation of our findings lies in a mixture of the hypothesis that the distribution of salient features differs between valences and the hypothesis that attention is captured more strongly by salient features of negative than positive or neutral stimuli.

Distribution of attention during retrieval predicts memory regardless of valence

Attention was only differentially allocated according to the accuracy and vividness of the memory during encoding. During recognition, eye fixations were more clustered for 'remembered' pictures relative to 'familiar' pictures, especially in negative remembered pictures regardless of arousal. This finding supports the idea that during recognition, vivid recollection may be prompted by enhanced memory for the salient features of the pictures. Positive images were also sampled more often than negative or neutral stimuli during encoding. These findings are mostly consistent with Sharot et al's (2008) findings and in some situations extended their findings, such as the findings for positive stimuli. Most previous studies have examined eye movement patterns during recognition to reveal that the eyes contain a lot of implicit information about what is remembered. For instance, Ryan and colleagues (2000) found that eye movements reflect previous exposure to a stimulus even without conscious awareness that the stimulus has been seen before. A neuroimaging study has taken advantage of this eye-movement-based memory effect to explore its relation to activity in memory related structures, i.e. the hippocampus (Hannula & Ranganath, 2009; reviewed in Kumaran & Wagner, 2009). As methods and design advance, future neuroimaging studies could take advantage of both pupillometry and eye-movement data to determine the neural correlates to the allocation of attention for positive, negative, and neutral information.

Limitations and Future Directions

Although our findings suggest that eye movements and pupillometry are both reliable indicators of subsequent memory and recollective experience during encoding and retrieval, these results do not speak to causation. In the case of eye movements, other factors such as distinctiveness might lead to clustering of fixations and enhancement of subsequent memory. Relatedly, pictures might not have been as emotionally arousing or distinctive because they were grey scale rather than color. Although there is evidence that arousal ratings of IAPS pictures do not differ between color and grey scale images (Bradley, Codispoti, Cuthbert, & Lang, 2001). Unfortunately, we cannot rule out whether the distinctiveness of emotional items was affected by the lack of color. In light of these caveats, we speculate that attention is allocated according to how salient features of a positive, negative, and neutral scene captures attention. As a result features are differentially encoded by valence through cognitive elaboration and subsequently recognized to produce an accurate and strong sense of remembering.

Pupil size differed as a function of emotional arousal with larger pupillary responses for emotionally arousing pictures after the initial light reflex during both encoding and retrieval. Another potentially interesting, but concerning finding was that pupil size differed on average as a function of emotional arousal during the initial light reflex at recognition. Particularly, this difference seems to appear just before end of the 1-second period after the onset of the stimulus (see Figure 4b). It is apparent that the trough of the change in pupillary response for positive and negative stimuli does not reach the same negative point before beginning to return to baseline. This difference was not found in the initial light reflex during encoding. Importantly, the same luminance-controlled stimuli were used during encoding and retrieval. Thus, this difference should not be attributed to any differences in the luminance of the stimuli divided by valence, which has been shown to modulate the magnitude of the initial light reflex (Bradley, Miccolo, Escrig, & Lang, 2008). The only difference between the 1st and 2nd session was the task being performed, incidental encoding and remember/know recognition. Although measures were taken to minimize the influence of factors, other than arousal, that affect the pupillary response, the pupillary response during retrieval might be contaminated with additional influences from the cognitive load of recognizing old and new items (Andreassi, 2000). We speculate that cognitive load and arousal during retrieval might act together to produce the differences in the light reflex and the faster return to baseline for positive and negative stimuli relative to neutral stimuli.

Another limitation was low proportions of freely recalled items. Because our main goal was to characterize the pupillary response and distribution of attention according to the participants' subsequent memory, optimizing the recognition task was given more priority. Thus, optimizing the delay for free recall was given less priority leading to low levels of free recall. Pilot experiments using a 24-hour delay revealed that recognition accuracy was too high to have the potential of dissociating reliable eye movement or pupillary response patterns for correctly remembered and later forgotten pictures. Free recall was also at floor levels with this 24-hour delay, so the decision to use a 7-day delay was made. A follow-up experiment using an immediate delay will be performed to properly assess whether stimuli that are later recalled show reliable pupillary response and eye movement patterns during encoding as was alluded to in the present experiments.

These findings have implications for future studies that could examine the interaction of MTL structures (amygdala and hippocampus) with regions involved in attention (prefrontal cortex, fusiform cortex) during encoding and retrieval of emotional memories. Previous studies have demonstrated that the lateral prefrontal cortex shows increased functional coupling with the hippocampus during retrieval of remembered or forgotten relational memories for neutral stimuli (Hannula & Ranganath, 2009). Taken with other results from their study, this finding suggests that memory retrieval may rely on the hippocampus, regardless of mnemonic accuracy, while interactions between the MTL and prefrontal cortex are necessary for accurate recollection of previously experienced relationships between stimuli (Kumaran & Wagner, 2009). Other studies have found that hippocampal amnesic patients do not attend to changed features of previously seen stimuli, while normal participants preferentially attend to the manipulated part of a stimulus even without subjective awareness of the change (Ryan, Althoff, Whitlow, Cohen, 2000). Although, there has also been evidence that awareness of the stimulus manipulation is necessary to see these hippocampal-dependent effects on eye movments (Smith, Hopkins, & Squire, 2006; Smith & Squire, 2008). The studies above particularly focus on relational memory and memory for previously seen but manipulated pictures. While relational memory was not the focus of the current study, it is conceivable that future studies could adapt the eye movement measures used in the current study to examine the functional activation and interaction of areas that mediate attention allocation to salient features of subsequently remembered stimuli. It would also be interesting to examine the extent to which the pupillary response, under controlled conditions of luminance and emotional content, correlates with amygdala activity or amygdala/hippocampal interactions. Broader future applications of this work include establishment of eye tracking and pupillometry as a fruitful approach to exploring hippocampal and amygdala function in settings where memory judgments and subjective arousal measures are difficult to obtain (e.g., infants; relational memory tests demonstrated in Richmond and Nelson, 2009; adults with mild cognitive impairment; non-human animals; as suggested in Kumaran & Wagner, 2009).

Conclusions

To our knowledge, this study is the first to examine whether physiological arousal as measured by the pupillary response is associated with subsequent memory during both encoding and retrieval. In addition, we examined whether attentional resources are differentially distributed during encoding and retrieval as a function of emotion and subsequent memory. The results of this study suggest that pupillometry indexes emotional arousal and memory during encoding and retrieval. This result suggests that larger pupillary responses during encoding and retrieval are associated with vivid recollection. The finding that allocation of attention during encoding for items later vividly recollected is dependent upon valence and arousal supports and extends the cue-utilization hypothesis proposed by Easterbrook (1959). Allocation of attention during retrieval is narrowed for subsequently remembered stimuli regardless of valence, which suggests that during recognition vivid recollection may be prompted by enhanced memory for the salient features of the pictures. The current study established that pupillometry and eye-movements are sensitive measures of various important psychological constructs in memory research, like attention and emotional valence and arousal. Overall, our findings support the claim that the pupillary response and allocation of attention during encoding and retrieval is indicative of how a stimulus is subsequently remembered.

References

Andreassi, J.L. (2000). Psychophysiology. LEA. Mahwah, NJ.

Bauer, P.J. (2008). Toward a neuro-developmental account of the development of declarative memory. *Developmental Psychobiology, 50*, 19-31. Ddoi: 10.1002/dev.20265

Bradley, M.M., Codispoti, M., Cuthbert, B.N., & Lang, P.J. (2001). Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion, 1*, 276-298.

Bradley, M. M., Miccolo, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology, 45*, 602-607.

Buchanan T.W. (2007). Retrieval of emotional memories. *Psychonomic Bulletin 133*, 761–779.

Canli, T., Zhao, Z., Brewer, J., Gabrieli, J.D.E., & Cahill, L. (2000). Event-related activation in the human amygdala associates with later memory for individual emotional experience. *J. Neuroscience, 20*(19), 5.

Canli, T., Desmond, J.E., Zhao, Z., & Gabrieli, J.D. (2002). Sex differences in the neural basis of emotional memories. *PNAS, 99*(16), 10789–10794.

Curran, T. (2004). Effects of attention and confidence on the hypothesized ERP correlates of recollection and familiarity. *Neuropsychologia, 42*, 1088–106.

Daselaar, S.M., Rice, H.J., Greenberg, D.L., Cabeza, R., LaBar, K.S., & Rubin, D.C. (2008). The spatiotemporal dynamics of autobiographical memory: Neural correlates of recall, emotional intensity, and reliving. *Cerebral*

Cortex, 18, 217-229. doi: 10.1093/cercor/bhm048

Dewhurst, S.A., & Conway, M.A. (1994). Pictures, images, and recollective experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 1-11.

du Bartas, G. (1578). Création du monde: First week, sixth day.

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

Fredrickson, B.L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist,* 56, 218–226.

Gable, P.A. & Harmon-Jones, E. (2008). Approach-related positive affect reduces breadth of attention. *Psychological Science,* 19(5), 476-482.

Gardiner, J.M., Gregg, V.H., & Karayianni, I. (2006). Recognition memory and awareness: occurrence of perceptual effects in remembering or in knowing depends on conscious resources at encoding, but not at retrieval. *Memory and Cognition, 34*, 227–39.

Hamann, S.B., Ely, T.D., Grafton, S.T., & Kilts, C.D. (1999). Amygdala activity related to enhanced memory for pleasant and aversive stimuli. *Nature Neuroscience, 2*, 289-293.

Hamann, S. (2001). Cognitive and neural mechanisms of emotional memory. *TICS*. 5(9), 394-400.

Hannula, D.E., Ranganath, C. (2009). The eyes have it: Hippocampal activity predicts expression of memory in eye movements. *Neuron*, 63(5), 592-599.

Hess, E. H., & Polt, J. M. (1960). Pupil size as related to interest value of visual stimuli. *Science*, *132*, 349–350.

Heuer, E, & Reisberg, D. (1992). Emotion, arousal, and memory for detail. In S.A. Christianson (Eds.), The handbook of emotion and memory: Research and theory (pp. 151-180). Hillsdale, NJ: Erlbaum.

Kensinger, E.A. (2004). Remembering emotional experiences: The contribution of valence and arousal. *Rev. Neurosciences*, 241–253.

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

Kensinger, E.A. (2009). Remembering the Details: Effects of Emotion. *Emotion Review, 1*(2), 99-113.

Kramer, T.H.I., Buckhout, R.I., & Eugenio, P.I. (1990). Weapon focus, arousal, and eyewitness memory: Attention must be paid. *Law and Human Behavior, 14*, 167-184.

Kumaran, D., Wagner, A.D. (2009). It's in my eyes, but it doesn't look that way to me. *Neuron, 63*(5), 561-563.

Labar, K.S. & Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nat Rev Neurosci, 7* (1), 54-64.

Laeng, B., Waterloo, K., Johnsen, S.H., Bakke, J., Låg, T., Simonsen, S.S., & Høgsæt, J. (2007). The eyes remember it: Oculography and pupillometry during recollection in three amnesic patients. *JOCN, 19*, 1888-1904.

Lang, P.J., Greenwald, M.K., Bradley, M.M., Hamm, A.O. (1993). Looking at pictures: affective, facial, visceral, and behavioral reactions.

Psychophysiology, 30, 261-73.

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

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). International affective picture system (IAPS): Affective ratings of pictures and instruction manual. *Technical Report A-8*. University of Florida, Gainesville, FL.

Libby, W. L., Lacey, B. C., & Lacey, J. I. (1973). Pupillary and cardiac activity during visual attention. *Psychophysiology, 10*, 270–294.

Loftus, E.F., Loftus, G.R., & Messo, J. (1987). Some facts about "Weapon Focus." *Law and Human Behavior, 11,* 55-62.

Mangels, J.A., Picton, T.W., Craik, F.I. (2001). Attention and successful episodic encoding: an event-related potential study. *Cogn Brain Res, 1*, 77–95.

McGaugh, J.L. (2000). Memory--a Century of Consolidation. *Science, 287*(5451), 248-251

Mickley, K.R., & Kensinger, E.A. (2008). Emotional valence influences the neural correlates associated with remembering and knowing. *Cognitive, Affective, & Behavioral Neuroscience, 8*(2), 143-152.

Ochsner, K. N. (2000). Are affective events richly recollected or simply familiar? The experience and process of recognizing feelings past. *Journal of Experimental Psychology: General, 129,* 242-261.

Paller, K.A., Wagner, A.D. (2002). Observing the transformation of experience into memory. *TICS, 6*, 93–102.

Pessoa, L., McKenna, M., Gutierrez, E. & Ungerleider, L. G. (2002). Neural processing of emotional faces requires attention. *PNAS, 99*, 11458–11463.

Phelps, E.A. (2004). Human emotion and memory: Interactions of the amygdala and hippocampal complex. *Current Opinion in Neurobiology, 14*, 198-202.

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

Richmond, J., and Nelson, C.A. (2009). Relational memory during infancy: evidence from eye tracking. *Dev. Sci. 12*, 549–556.

Rowe, G., Hirsh, J.B., & Anderson, A.K. (2006). Positive affect increases the breadth of attentional selection. *PNAS, 104*, 383–388.

Ryan, J.D., Althoff, R.R., Whitlow, S., Cohen, N.J. (2000). Amnesia is a deficit in relational memory. *Psychological Science, 19*, 454–61.

Sharot, T., Delgado, M.R., Phelps, E.A. (2004). How emotion enhances the feeling of remembering. *Nature Neuroscience, 7*, 1376–1380.

Sharot, T., Davidson, M. L., Carson, M. M., & Phelps, E. A. (2008). Eye Movements Predict Recollective Experience. *PLoS ONE , 3*, 2884. doi: 10.1371/journal.pone.0002884.t002

Smith, A.P., Stephan, K.E., Rugg, M.D., & Dolan, R.J. (2006). Task and content modulate amygdala-hippocampal connectivity in emotional retrieval. *Neuron, 49*, 631–638

Smith, C.N. & Squire, L. (2008). Experience-dependent eye movements reflect hippocampus-dependent (aware) memory. *J. Neuroscience*. 28(48), 12825-12833.

Steblay, N. M. (1992). A meta-analytic review of the weapon focus effect. *Law* *and Human Behavior, 16*, 413–424.

Tulving, E. (1983). Elements of Episodic Memory. London: Oxford Univ. Press.

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

Vuilleumier, P., Richardson, M. P., Armony, J. L., Driver, J. & Dolan, R. J. (2004). Distant influences of amygdala lesion on visual cortical activation during emotional face processing. *Nature Neuroscience, 7*, 1271–1278.

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

Yonelinas, A.P. & Jacoby, L.L. (1994). Dissociations of processes in recognition memory: effects of interference and of response speed. *Canadian Journal of Experimental Psychology*, 4, 516–35.

Footnotes

[1] The picture size on the screen was about 2 times larger than the pictures in Sharot et al. (2008; 5.5 x 5.5 in), thus the average inter-fixation distance was proportionally larger when comparing this study and Sharot's measurements in a standard measurement format (inches).

Table 1

Proportion of overall recognition and recall for positive, negative, and neutral items

	Hit		FA		Misses		Corr. Rej.		Recalled	
Memory Type	*M*	*SEM*	*M*	*SEM*	*M*	*SEM*	*M*	*SEM*	*M*	*SEM*
Overall	0.74	0.03	0.23	0.03	0.26	0.03	0.77	0.03	0.02	0.00
Positive	0.74	0.03	0.23	0.04	0.26	0.03	0.77	0.04	0.02	0.00
Negative	0.81	0.02	0.28	0.03	0.19	0.02	0.73	0.03	0.04	0.01
Neutral	0.66	0.04	0.18	0.02	0.34	0.04	0.82	0.02	0.01	0.00

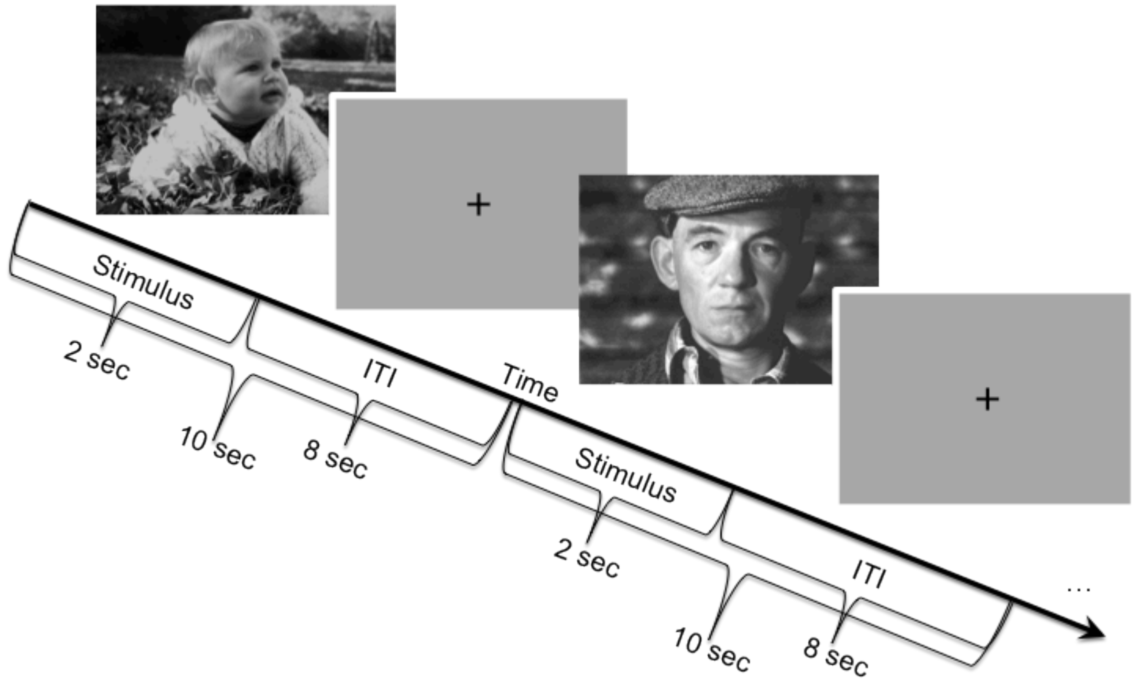

Figure 1a. Representative encoding session trial.

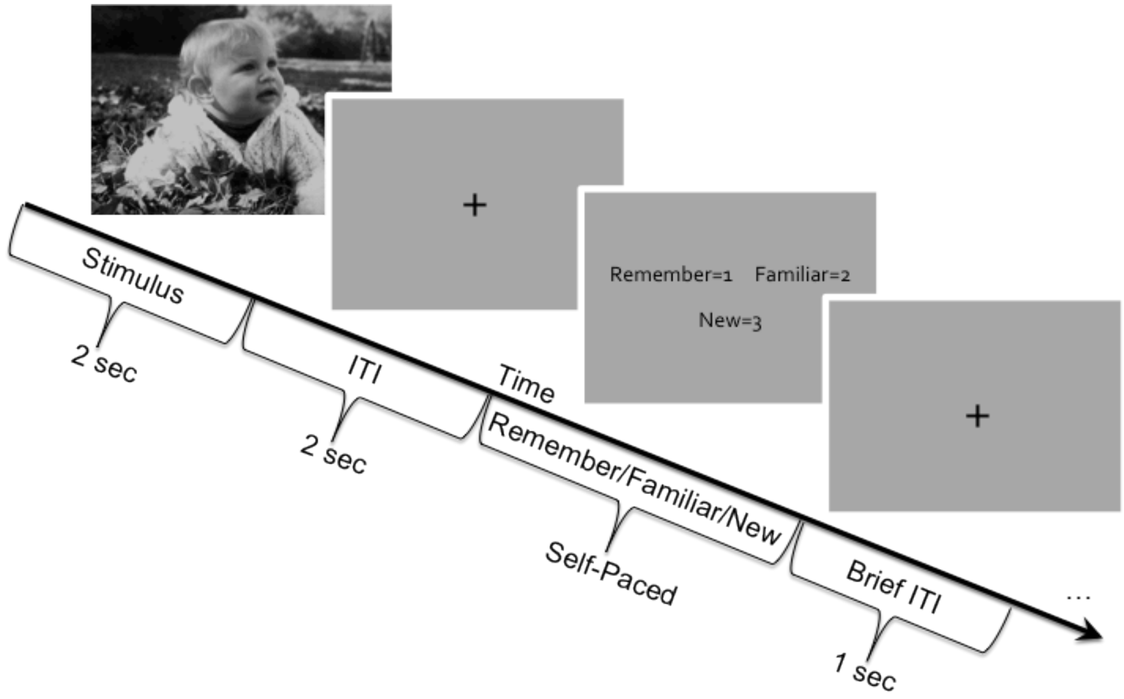

Figure 1b. Representative recognition session trial.

Figure 2. Eye-gaze pattern. The average inter-fixation distance was calculated by summing the distance between two fixation coordinates in pixels (red lines) and dividing that by the number of fixations over a given stimulus presentation period. The fixation frequency is the sum of number of fixations over a given stimulus presentation period. The diameter of each circle depicts the relative duration of each fixation.

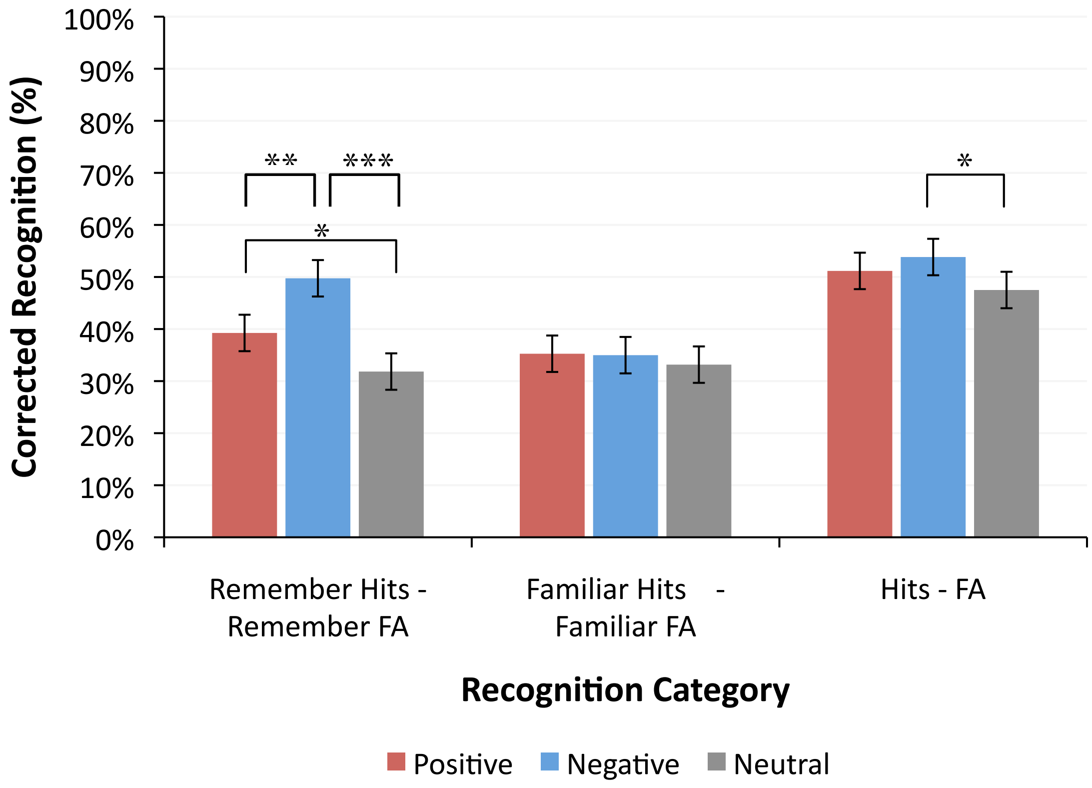

Figure 3. Corrected recognition by valence for overall recognition and recognition split by subjective experience (remember and familiar). Error bars depict the SEM for each category. *** $p<0.0005$, **$p<0.005$, *$p<0.05$

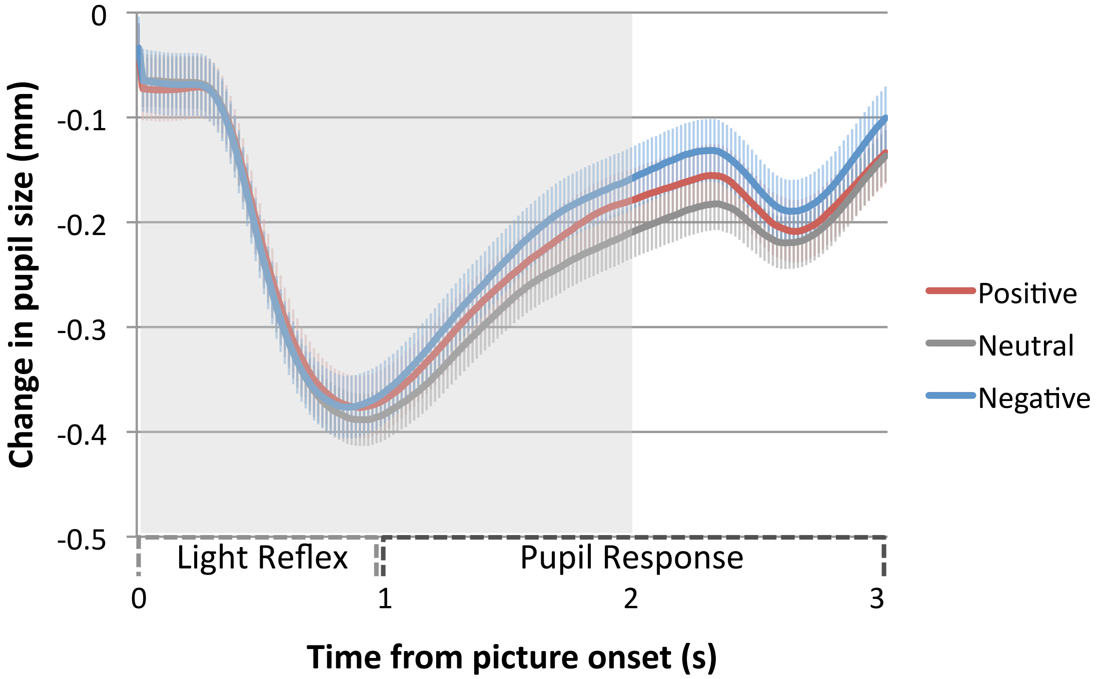

Figure 4a. Average pupil response split by emotional valence during encoding. The gray shaded area depicts the stimulus presentation period. The dashed vertical lines depict the windows for the light reflex (0-1 sec) and the pupil response (1-3 sec). The colored shading area around each valence line is the SEM for positive, negative, and neutral, respectively. Faster return to baseline shows more dilation or less average constriction.

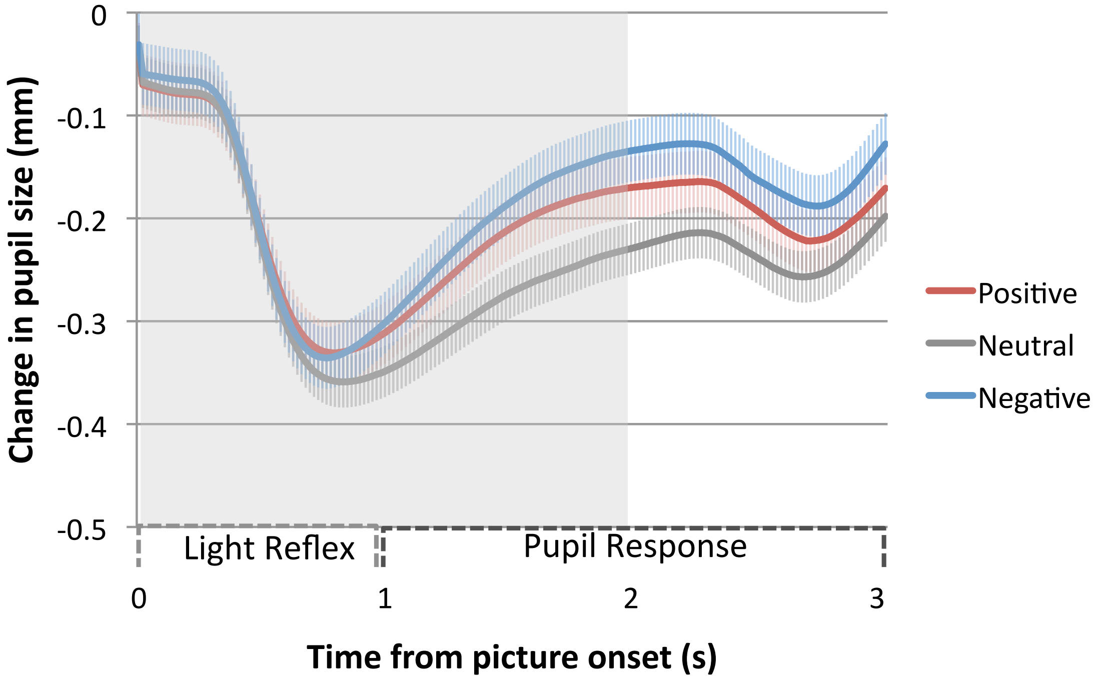

Figure 4b. Average pupil response split by emotional valence during retrieval. The gray shaded area depicts the stimulus presentation period. The dashed vertical lines depict the window for the light reflex (0-1 sec) and the pupil response (1-3 sec). The colored shading area around each valence line is the average SEM of the mean over the presentation period for positive, negative, and neutral, respectively. Faster return to baseline shows more dilation or less average constriction.

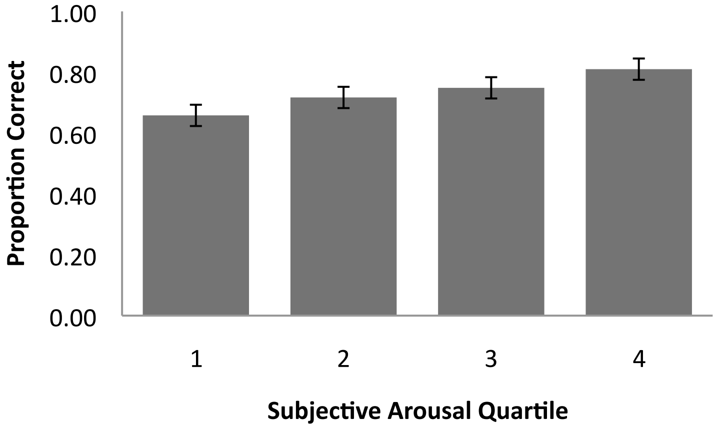

Figure 5a. Proportion of items correctly recognized as a function of subjective arousal quartiles (1= low arousal quartile, 4=high arousal quartile). Error bars depict the SEM.

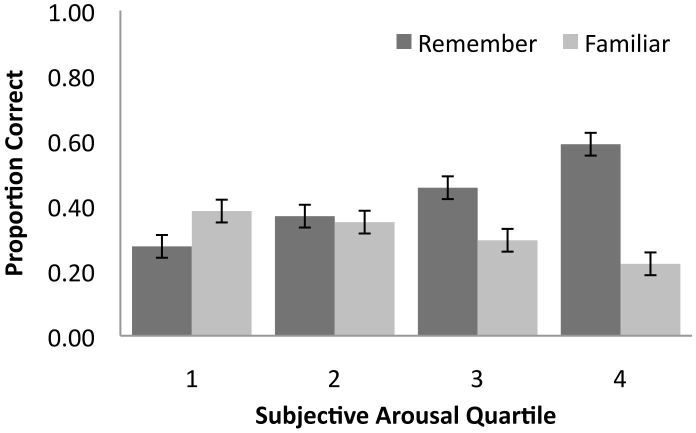

Figure 5b. Proportion correct remember and familiar judgments as a function of subjective arousal quartiles (1= low arousal quartile, 4=high arousal quartile) and recollective experience (remember or familiar). Error bars depict the SEM.

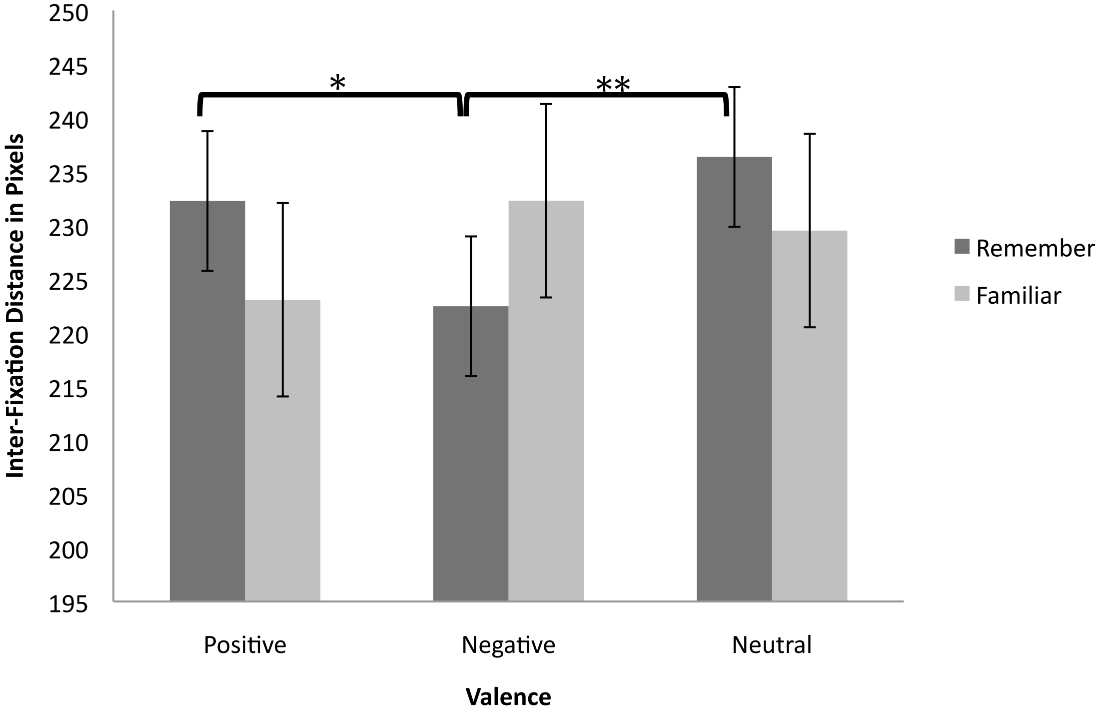

Figure 6a. Average inter-fixation distance as a function of valence and recollective experience during encoding. Error bars depict the SEM. This figure depicts more clustered fixations during encoding for negative recollected items than positive or neutral recollected items. $^{**}p<0.005$, $^{*}p<0.05$

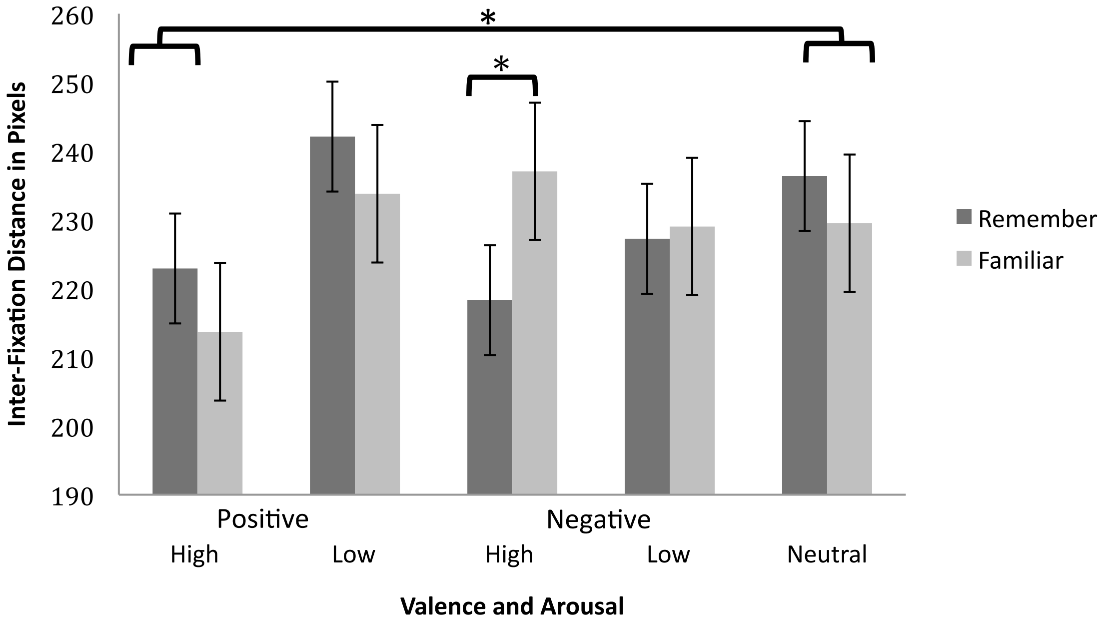

Figure 6b. Average inter-fixation distance during encoding as a function valence, physiological pupillary arousal, and recollective experience. Error bars depict the SEM. Note that inter-fixation distance is characterized by a main effect of valence with more clustering for high arousal positive and negative items relative to neutral items. In addition, note that we find more clustering for high arousal negative items that were subsequently recollected than high arousal negative items that were subsequently familiar. *p<0.05

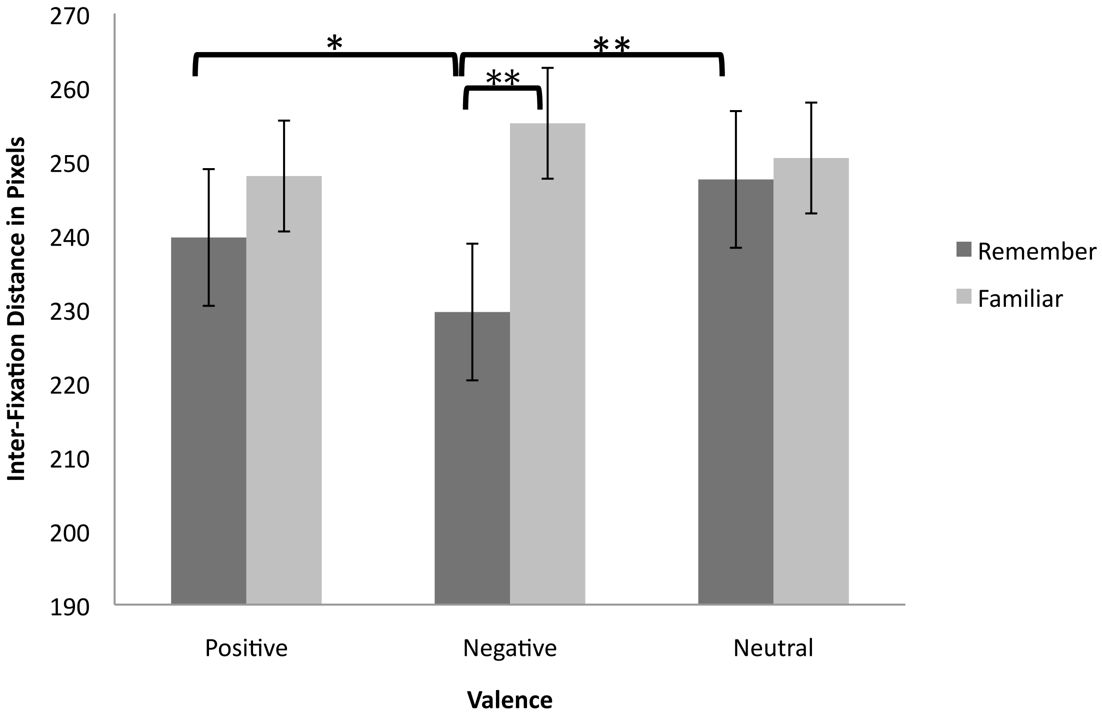

Figure 7a. Average inter-fixation distance as a function of valence and recollective experience during retrieval. Error bars depict the SEM. This figure depicts more clustered fixations for negative items which were subsequently recollected than negative items which were subsequently familiar *** $p<0.0005$, **$p<0.005$, *$p<0.05$

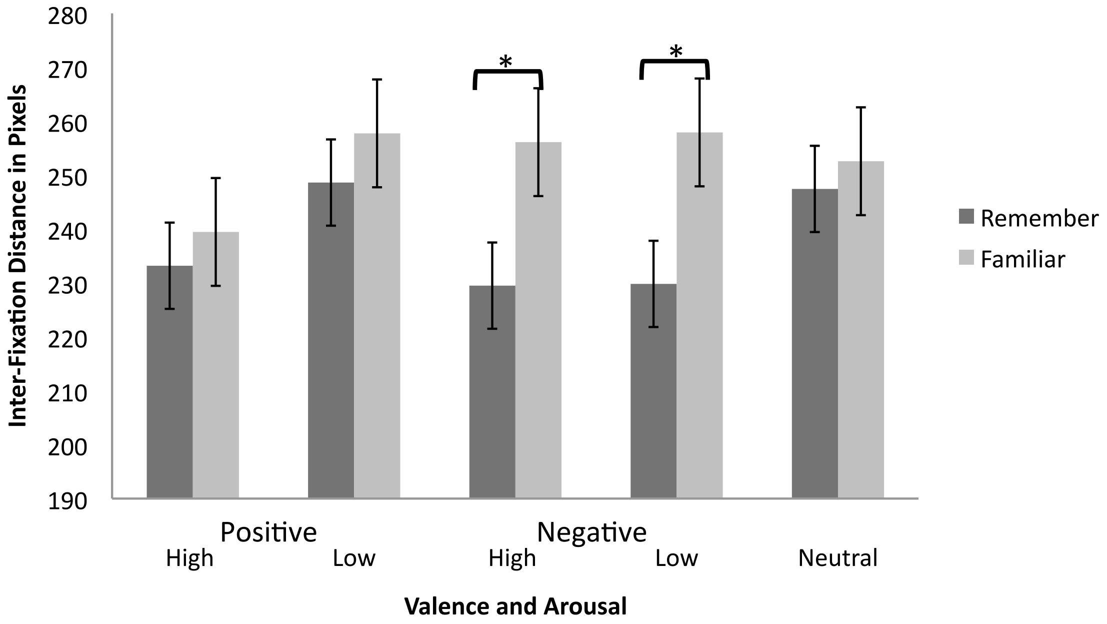

Figure 7b. Average inter-fixation distance during retrieval as a function valence, physiological pupillary arousal, and recollective experience. Error bars depict the SEM. Consistent with inter-fixation distance during encoding, this figure depicts more clustering for negative items that were subsequently recollected thannegative items that were subsequently familiar, regardless. $*p<0.05$

Figure 8. Example positive, negative, and neutral item fixation paths during encoding for vividly recollected items. This figure depicts examples of a less clustered average inter-fixation distance (IFD) for positive (top left; IFD=436 pixels) and neutral (bottom; IFD=325) items and a more clustered average inter-fixation distance for negative (top right; IFD=159) items.

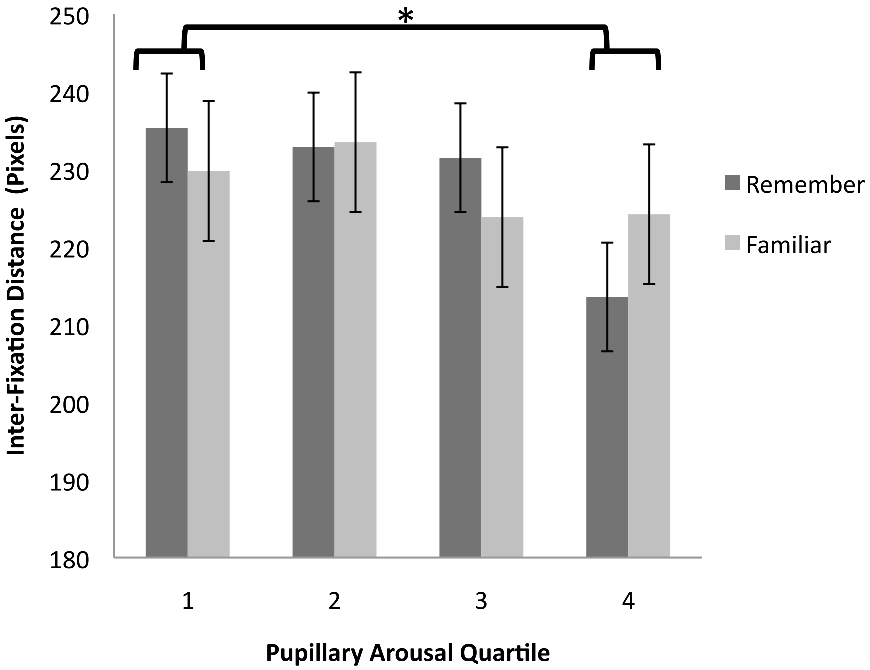

Figure 9. Average inter-fixation distance during encoding as a function of physiological arousal, measured by pupil response, and subsequent recollective experience during encoding (1= low arousal quartile, 4=high arousal quartile). Error bars depict the SEM. This figure depicts more clustered fixations for highly physiologically arousing items than minimally physiologically arousing items. *$p<0.05$