RESEARCH ARTICLE

Memory for non-painful auditory items is influenced by whether they are experienced in a context involving painful electrical stimulation

Keith M. Vogt[1,2] · Caroline M. Norton[1] · Lauren E. Speer[3] · Joshua J. Tremel[4] · James W. Ibinson[1] · Lynne M. Reder[2,5] · Julie A. Fiez[2,3,4]

Abstract

In this study, we sought to examine the effect of experimentally induced somatic pain on memory. Subjects heard a series of words and made categorization decisions in two different conditions. One condition included painful shocks administered just after presentation of some of the words; the other condition involved no shocks. For the condition that included painful stimulations, every other word was followed by a shock, and subjects were informed to expect this pattern. Word lists were repeated three times within each condition in randomized order, with different category judgments but consistent pain-word pairings. After a brief delay, recognition memory was assessed. Non-pain words from the pain condition were less strongly encoded than non-pain words from the completely pain-free condition. Recognition of pain-paired words was not significantly different than either subgroup of non-pain words. An important accompanying finding is that response times to repeated experimental items were slower for non-pain words from the pain condition, compared to non-pain words from the completely pain-free condition. This demonstrates that the effect of pain on memory may generalize to non-pain items experienced in the same experimental context.

Keywords Pain · Memory · Remember-know · Electric nerve stimulation · Recollection · Familiarity

Introduction

The experience of acute somatic pain is often memorable, the quintessential example being the lesson learned from touching a hot stove. Painful stimulation is sufficient to

✉ Keith M. Vogt
  Keith.Vogt@Pitt.edu

1 Department of Anesthesiology and Perioperative Medicine, University of Pittsburgh, School of Medicine, Pittsburgh, PA, USA

2 Center for the Neural Basis of Cognition, Pittsburgh, PA, USA

3 Department of Neuroscience, University of Pittsburgh, Pittsburgh, PA, USA

4 Department of Psychology, University of Pittsburgh, Pittsburgh, PA, USA

5 Department of Psychology, Carnegie Mellon University, Pittsburgh, PA, USA

engage the hippocampus and amygdala (Bingel et al. 2002), explaining not only memory formation, but also why one may have physiologic and emotional responses to pain. Aside from the memory of pain itself, noxious stimuli can affect learning and memory for other information, such as the stimuli associated with painful outcomes or experiences. How pain influences different aspects of learning and different types of memory is nuanced, and the body of related research is broad. To our knowledge, there has been little human research that closely examines the impact of acute pain on long-term memory of an event, with conflicting results across studies (described below). There is a literature of related research that we briefly review below, including the effects of emotion on memory and pain on working memory.

Recent reviews show an overall memory advantage for emotionally valanced experimental items (LaBar and Cabeza 2006; Yonelinas and Ritchey 2015). To briefly summarize the usual methodology, emotionally disturbing images are presented and recognition testing subsequently used to measure explicit memory (Yonelinas and Ritchey 2015). Though the field of emotional memory research does not

typically involve painful stimulation, there may be mechanistic overlap, as the experience of pain has varying unpleasantness ratings that could be construed as emotionally valanced (Rainville et al. 1997; Schreckenberger et al. 2005; Choi et al. 2006). If this is the case, words experienced in the context of pain should produce memory effects similar to those observed for emotional versus neutral items. For instance, recollection of specific contextual details is more affected than familiarity (non-specific recognition) (Yonelinas and Ritchey 2015); though see (Bennion et al. 2013) for conflicting results. There may also be lingering effects of emotional arousal, as encoding of neutral items after a period of emotional arousal can also show improved recollection (Lonsdorf et al. 2017). Additionally, consolidation and memory decay may play a role, as differences due to emotional valence are better detected after 24-h, compared to immediately (Yonelinas and Ritchey 2015).

Another relevant area of research focuses on pain's effect on working memory tasks, with performance typically inhibited by experimental pain stimuli (Hood et al. 2013) and chronic pain (Berryman et al. 2013). Such work demonstrates that pain can modulate attentional resources, but there is a clear distinction between working memory and long-term memory formation. It is accepted that working memory and long-term memory recruit distinct cerebral resources, with working memory experimental tasks involving prefrontal cortex (D'Esposito 2007) while successful encoding into long-term memory represents an additional hippocampal-predominant process (Wais 2008) influenced by attention and autonomic arousal (Mather and Sutherland 2011). Though maintenance in the working memory buffer is a necessary step prior to long-term storage, poorer performance on a task assessing working memory performance can be inversely correlated to incidental encoding into long-term memory, assessed by recognition testing (Zanto et al. 2016).

A related consideration is that attentional arousal can increase the bias toward memory for perceptually important stimuli (Mather and Sutherland 2011). This advantage comes at the expense of less relevant stimuli, and depends on the attentional division between the arousing and task stimuli (Mather and Sutherland 2011). Similar parallels can be drawn from the motivation literature, in which threatening stimuli can impair hippocampal-driven binding processes (Murty and Adcock 2017). Acute stress during encoding can either enhance or impair episodic memory, depending on similarity and timing relative to the materials to be encoded (Shields et al. 2017). Painful stimulation might be inferred to be similar (or dissimilar) to any of these paradigms, as there is certainly conceptual overlap. However, the direct evidence that is needed to draw such comparisons is lacking.

The present study directly tests for an effect of pain on human explicit memory. It builds upon previous studies that have examined recognition memory for items experienced in the context of pain. In one study (Forkmann et al. 2016), a surprise recognition memory test was administered several minutes after participants viewed pictures that were or were not paired with a painful thermal stimulation. Memory was worse for pictures previously paired with pain, compared to those not paired (Forkmann et al. 2016). In a similar image-categorization paradigm, association with painful electric shocks had no effect on immediate recognition memory testing, suggesting no effect of pain on memory (Schwarze et al. 2012). However, in a second cohort who underwent the same encoding experience but who were tested the following day, memory performance was enhanced for pain-paired items (Schwarze et al. 2012). These disparate results have several possible explanations, including the obvious experimental design differences. In addition to different intervals between encoding and memory testing, the type, duration, consistency of pairing, and timing of the pain stimulus vary across experiments, all of which may contribute to divergent results. These subtle design differences highlight why more confirmatory work is needed to discern which specific experimental factors consistently modulate the influence of pain on memory formation.

The present study shares similarities with these past works. Like previous investigations, the effect of acute experimental pain on subsequent recognition memory is assessed. One distinctive feature of our work is that spoken words, rather than visual images, were used as stimuli. The primary experimental manipulation, done during encoding, allows for comparison between items that were pain-paired and two types of non-pain items that differ in their proximal relationship to the pain events. Our primary hypotheses were designed around these two fundamental comparisons. We first predicted that pain would impair memory performance; pain-paired items would have lower recollection than non-pain items that were experienced within the same experimental context. This hypothesis was motivated by prior studies demonstrating that pain-paired items have decreased hippocampal activation (Forkmann et al. 2013). We further expected that an isolated improvement in familiarity responses might be seen with pain-paired items, consistent with previous work (Schwarze et al. 2012). Second, we anticipated that pain context would enhance recollection memory, due to increased overall arousal during the experimental condition containing painful stimulations. This was tested by comparing non-pain items between the pain-containing and pain-free portions of the experiment, with the prediction that non-pain words interleaved with pain-paired words would show better recognition than non-pain words experienced in an entirely pain-free experimental epoch.

Because of the pleotropic behavioral effects of pain mentioned above, we also sought to detect implicit measures of memory influenced by pain-pairing and pain context. By

implicit memory, we mean behavioral effects for which the subject is unaware. For example, repetition priming tends to produce reductions in response times to stimuli even when subjects cannot recollect having seen the stimulus previously (Henke 2010; Squire and Dede 2015). This performance enhancement, compared to response times for novel items is subconscious and best known for being hippocampal independent, particularly from lesion studies (Cave and Squire 1992). Priming effects can also be used to demonstrate differences in learning between experimental conditions, revealing meaningful behavioral differences that occur through non-hippocampal mechanisms (Francis 2014).

Another measure of implicit memory for acute pain is learned physiologic response, mostly explored in fear (or threat) conditioning research. The typical design of a conditioning experiment includes unconditioned (typically painful) stimuli, paired to some set of neutral stimuli (LeDoux 2014; Lonsdorf et al. 2017). For items previously associated with noxious stimuli, re-exposure to a conditioned stimulus causes a sympathetic response, which can be measured by changes in heart rate (Castegnetti et al. 2016) or electrodermal activity (Schiller et al. 2010). The magnitude of the conditioned response at a later time point is used to determine the strength of the underlying memory trace (Dunsmoor and Kroes 2019). Interestingly, the sympathetic response from this learned association may persist up to 1 year later (Schiller et al. 2010), indicating the possibility of a prolonged aversive association for items previously experienced with pain. In an experimental design referred to as category conditioning (Dunsmoor and Kroes 2019), episodic memory enhancement is seen for CS + items; in partial-reinforcement paradigms, this improvement persists for items not directly paired with the US. Another interesting aspect of this research examines the extent to which a conditioned response to one stimulus generalizes to another. Subjects' explicit ratings of negativity or anxiety do not seem to generalize across stimuli, while involuntary startle responses do generalize (Andreatta et al. 2015). Subjects that are not explicitly aware of the relationship between conditioned and unconditioned (painful) stimuli have generalized electrodermal responses to the experimental context up to 1 month later, and exhibit more avoidance behavior (Grillon 2002).

We examined two possible measures of implicit memory as secondary outcomes. The first measure was repetition priming, reflected by the change in average response time during the encoding task. Decreased response time with repeated exposure to the same stimuli is an expected effect of practice, but we hypothesized that pain-paired words would show a blunting of this facilitation. This prediction was based on the expectation that concomitant pain stimulation would distract attention and impair the performance improvement (Keane et al. 2015) that would otherwise shorten response time when making a different decision

about a word heard previously in the experiment. We also monitored electrodermal activity and heart rate, to assess for sympathetic changes seen with repeated exposure to words during both the encoding task and during memory testing (when no pain stimuli are employed). We hypothesized that these physiologic responses would not show generalization; only pain-paired words were expected to trigger a sympathetic response, as subjects were anticipated to be explicitly aware of the pain-pairings.

Methods

Subjects

Data were acquired from healthy volunteer subjects between the ages of 18 and 30 who were recruited from the university community. The cohort analyzed consists of 30 subjects (22 female) with age in years 23.3 ± 2.8 (mean $\pm$ standard deviation). They received either course credit or \$10 per hour as compensation. Eligibility was determined by self-report of exclusionary criteria and also by performance on a brief memory task during a screening visit. All subjects were free from significant memory impairment, hearing loss, sleep apnea, chronic pain, other chronic medical problems, neurologic and psychiatric diseases, as well as the use of antidepressants, antipsychotics, antihistamines, antianxiety medication, stimulants, sleep aids, and pain medication. The study was approved by the University of Pittsburgh Institutional Review Board (PRO16110197) and conforms to all relevant standards for the ethical and responsible conduct of research.

Procedures

Screening procedures

Subjects came into the testing location for two visits separated by at least 1 day. The first visit involved a paper and pencil screening to determine eligibility and to obtain responses to a series of psychometric questionnaires, including an assessment of memory ability. The subjects completed a written consent form specific to the screening procedures. The brief memory screening task consisted of a series of 30 word-pairs printed on a sheet of paper. Subjects marked whether they thought the words in each pair were related or unrelated, and were not informed they would later be tested for memory. Subjects then completed four psychometric questionnaires: State-Trait Inventory for Cognitive and Somatic Anxiety (STICA) (Gros et al. 2007), Pain Anxiety Symptom Scale—Short Form 20 (PASS) (McCracken and Dhingra 2002), Pain Vigilance and Awareness Questionnaire (PVAQ) (Roelofs et al. 2003), and Pain

Catastrophizing Scale (PCS) (Osman et al. 1997). After spending several minutes on these questionnaires, subjects did a memory test in which all words seen previously were printed, with half of the words paired differently. Subjects were asked to indicate whether the words had previously been seen together as a pair or were previously paired with different words. These data were used as a screening test for memory ability, and not further analyzed in the experiment. Subjects whose d' (calculation described below) was less than 0.5 on the memory screen would have been excluded from the main experiment. However, no screened subjects were below this threshold.

Learning task procedure and design

On arrival for the second visit, subjects completed a separate written informed consent process for the main part of the experiment, including consent for painful nerve stimulation. They were then oriented to the experiment and the nerve stimulator. The amplitude of stimulation was titrated in 1 mA increments to a targeted subjective rating of 7

out of 10 pain, using a numerical rating scale in which 0 was anchored to no pain and 10 to the worst pain imaginable. Prior to the learning portion, subjects went through a brief practice session in which they heard six words and made each of the six categorization decisions used subsequently in the experiment. One of the words in the practice session was heard twice and was followed each time by a 1.5-s electric shock. After the practice session, a pain rating was again obtained, and the nerve stimulator intensity adjusted, if necessary, to achieve a 7/10 rating. After this second adjustment, the intensity of the stimulation was not further manipulated during the subsequent experiment, and the current flow (which was maintained at a constant value for all stimulations) was recorded. A second pain rating was obtained immediately after the pain condition of the experiment; pain ratings for individual items were not obtained.

The design of the main part of the experiment is depicted graphically in Fig. 1. There were two learning conditions: the Pain condition during which every other word was paired with a painful electric shock and the No Pain condition, during which no pain was experienced for any word.
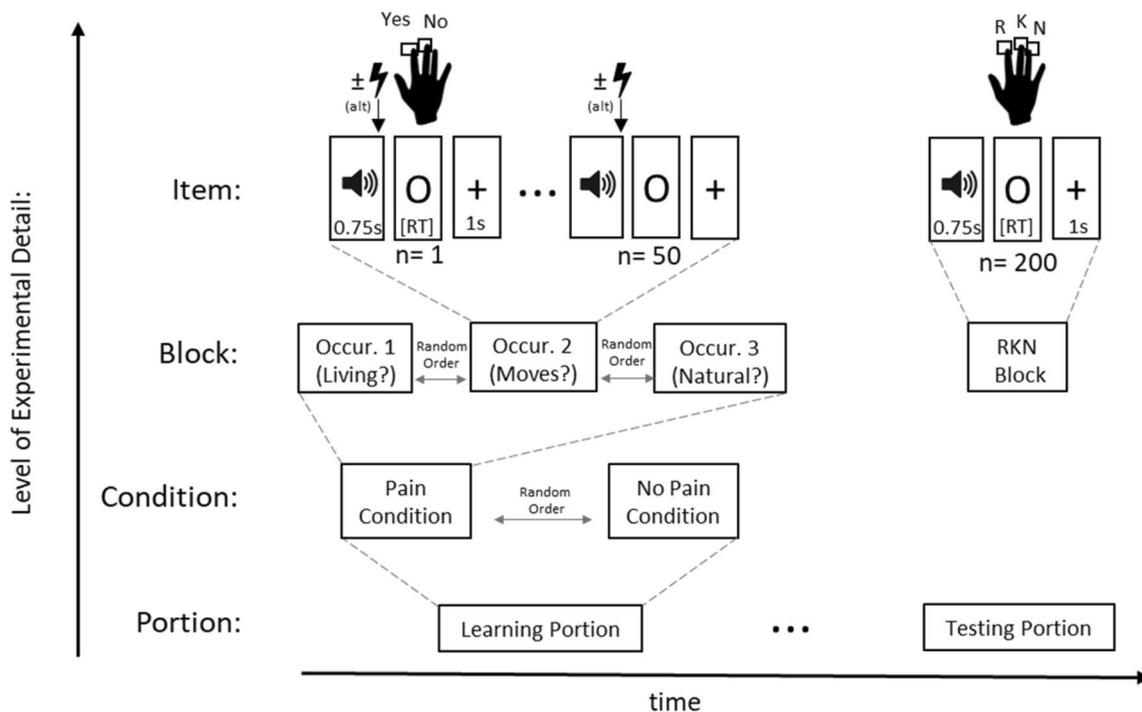

Fig. 1 Graphical depiction of experimental design in nested layers of detail, labeled on the left. Time is represented along the *x* axis. Increasing level of experimental detail is shown going from bottom to top, with dashed lines indicating expanded detail of items within a larger hierarchical part of the experiment. The two main portions of the experiment, Learning and Testing, are shown as blocks at the bottom, these always occurred consecutively. The Learning portion has two conditions: Pain and No Pain, and the order of these was randomized. Each Learning condition consisted of three repetitions of a word list, examples of which are shown expanded above the Pain condition. Both Pain and No Pain conditions contained three occurrences of the same 50-word list. As shown in the item level of detail, words were delivered auditorily, over 0.75 s. In the Pain Condition (only) a painful electric shock immediately followed every other word (50% shock contingency). No shocks occurred in the No Pain condition or in the Testing portion of the experiment. *RT* reaction time, *RKN* remember, know, new, *Alt* alternating

Each condition consisted of a list of 50 words repeated 3 times, in random order. In describing the experiment and analysis, items from the learning portion of the experiment were categorized into three groups. Pain Mixed words were the 25 words paired with a painful shock in the Pain Condition. No Pain Mixed words were the 25 words in the Pain condition that were not pain-paired. No Pain Alone words were the 50 words experienced in the No Pain condition (which was completely without noxious stimulation). Words were presented aurally, one at a time, through headphones and subjects made decisions about them while seated at and interacting with a laptop computer. Subjects were to judge each word on a particular dimension for that list. There were six different judgment categories, three assigned randomly to each of the two learning conditions. The same words were repeated in different random orders within the two learning conditions. When shocks did occur, the onset was immediately after the end of the word and it was 1.5 s in duration. Subjects were able to respond any time after the start of the word being played, as well as during the subsequent shock, if one occurred. The maximum response time (RT) window was 6 s, after which the next word would be presented automatically. Within each repetition, subjects made one of six categorization decisions about the word: (1) whether it moves or not, (2) whether it is living or not, (3) whether it is natural or not, (4) whether it would fit in a shoebox or not, (5) whether it is a place or event or not, and (6) whether it is typically used or not. Subjects responded by pressing a key with their index finger to indicate a "yes" response or with their middle finger to indicate a "no" response. Words in the pool were such that approximately 50% had yes and no responses for each categorization question. However, because of similarities in concept properties, this was accomplished with the moves, living, and natural decisions grouped together and used in one condition, while the shoebox, place, and use decisions were used in the other. The order of the Pain versus No Pain conditions was randomized to occur first versus second across subjects. The assignment of a set of decisions to a condition, the order of the decision list within a condition, and the order of the words within a decision list were randomized for each subject, with the constraint that half of the subjects received the pain context first in the experiment. All subjects had 50 words assigned to each of the 6 possible categorization decision tasks.

Memory testing

Explicit memory testing occurred after the learning portion, following a break of several minutes for instructions. Recognition testing was employed, using the Remember-Know-New (RKN) scheme (Migo et al. 2012). Subjects were given a printed sheet describing the Remember, Know, and New responses for the RKN procedure and were asked to read

it. These instructions parallel those established by previous investigations (Rajaram 1993), and a copy is available as an Appendix to the supplementary materials. To ensure comprehension of the task, and maximize the distinction between Remember and Know responses, subjects were asked to explain the RKN procedure back to the investigator. A standardized recording summarizing these instructions was also played to the subject before beginning the RKN test. All the words previously heard in the learning portion of the experiment were then played, as well as an equal number of foils (200 items total). The order of the words was randomized, and the assignment of a word from the word pool to be a foil versus used in the experiment was also random. Physiologic monitoring was continued during the RKN portion of the experiment. No shocks were delivered, and subjects were informed of this.

Stimuli

Auditory stimuli were randomly drawn from a bank created for this experiment which contained 720 English nouns. Words were obtained from the English Lexicon Project (Balota et al. 2007) database (http://elexicon.wustl.edu/). The common logarithm of the Hyperspace Analogue to Language (HAL) frequency (Lund and Burgess 1996) was used to narrow words, using log(HAL) greater than six and less than or equal to 11 as inclusion criteria. The log(HAL) is a measure of how common a word is in the English language, so this was to ensure that words used in the experiment existed across a contiguous range of how often they appear in usage. Words were further constrained to contain between one and four syllables and between four and eight phonemes, such that each could be spoken in approximately the same amount of time. Words were then reviewed by one of the investigators to prune this list. Exclusion criteria included words that were proper nouns, the plural form of a word already included, vulgar or otherwise objectionable, pronounced similarly to a word already included, ambiguous based on pronunciation (e.g., red versus read), Greek letters, a slang or colloquial term, representative of an abbreviation, or starting with the same sub-word as a word already included. Words with possible pain relevance meeting these criteria were maintained in the pool, though these represent a small fraction of the total number of words available for use. Assignments of any word to a condition (or to be used for an individual subject at all) was randomized.

Digital audio recordings were made by one male member of the study team speaking each word. Recordings were captured and edited using the open source software, Audacity 2.1.0 (http://audacity.sourceforge.net/) and saved as mp3 files. For timing alignment, all sound files were adjusted so that the length was exactly 750 ms (ms) and the end of the word coincided with the end of the file.

Equipment

Pain stimuli were generated using an electric nerve stimulator (EZstim II, Life Tech, Inc, Norcross, GA, USA). Two electrodes were attached to the lateral aspect of the subject's left index finger, straddling the proximal interphalangeal joint. Subjects were then connected to physiologic monitoring equipment by applying common electrodes to their chest for electrocardiogram monitoring, though these data are not reported here due to technical issues. Electrodermal activity (EDA) was acquired from the left palm, with electrodes on the hypothenar and thenar eminences. The left palm was cleaned with distilled water and allowed to air dry prior to electrode application (EDA Isotonic Gel Electrodes, Item EL507, BIOPAC Systems, Goleta, CA, USA). Electrodes were also prepared by adding a small amount of Isotonic Recording Electrode Gel (Gel 101, BIOPAC Systems, Goleta, CA, USA). Data were digitized using a BIOPAC MP160 (BIOPAC Systems, Goleta, CA, USA) data acquisition unit and Acqknowledge version 5.0 (BIOPAC Systems, Goleta, CA, USA), running on a Windows 10 laptop PC.

All parts of the experiment were implemented with E-Prime version 2.0 (Psychology Software Tools, Sharpsburg, PA). The pairing of words with electric shocks was accomplished using custom hardware that allowed E-Prime to control relays via a serial-emulated USB connection. The nerve stimulator was modified so that the push-button switches on the front of the device could be closed electronically using computer control. Response time data were logged by E-Prime during the experiment, and these intervals were synchronized to begin at the onset of the word stimulus. A 5 V square-wave logic signal indicating the timing and type of each word being played was generated by a separate bank of relays, also controlled by E-Prime. These trigger signal waveforms were recorded using the BIOPAC and allowed for alignment of the experimental events with physiologic data recordings.

Data analyses

All statistical analyses were carried out in SPSS Statistics 23, using alpha = .05 to determine statistical significance, with Bonferroni correction for multiple comparisons applied whenever applicable. Histograms were generated with SPSS. Other data were organized, sorted, and displayed using Microsoft Excel 2016. Outlier removal was performed using RStudio (version 1.0.153, https://www.rstudio.com/) running R version 3.2.5. The median absolute deviation (MAD) was calculated as the absolute value of the difference between the median and each RT value (Leys et al. 2013). RT data with MAD > 3.5 were identified as outliers. Incorrect RKN responses were removed from RT analysis. For learning response times, the data were skewed, and

therefore, transformed prior to analysis. Several transformations were explored, including taking the square, square root, common logarithm (log_{10}), and natural logarithm. The common logarithm was the transformation found to move the data into a pattern adequately approximating a normal distribution. The RT analysis was performed on log-transformed data, using a repeated measures linear mixed model (LMM) analysis with a Bonferroni adjustment for multiple comparisons. Compound symmetry covariance structure was selected, to account for the within-subjects design. Model terms were main effects of trial and condition, as well as an interaction term between these two variables. Transformed RKN response times were also analyzed using an LMM, with response type (remember and know) and each different word-type as separate factors. Though all RT analyses were done in $log_{10}(RT)$ space, the RT data in the figures are shown in milliseconds for conceptual clarity.

Signal detection theory was used to estimate memory sensitivity, separate from response bias. To score memory screening results, hits were correct pair judgments, and false alarms were pairs incorrectly identified as previously paired. Using the formula z(hits)–z(false alarms), d' was calculated, where the function $z()$ represents the cumulative Gaussian distribution. Using a threshold of $d' > 0.5$ on the memory screen, no subjects were excluded based on memory performance. Memory performance for the main experiment was calculated similarly as a function of the hit and false alarm rates for experimental items. d' was calculated for the combined Remember and Know responses and Remember responses. The false alarm rate was calculated as the proportion of foils designated (incorrectly) with either a Remember or Know response. For experimental conditions for which a subject had a perfect hit rate, or no false alarms, the input to the d' formula (which can be neither one nor zero) was estimated as having one incorrect response in a field of twice as many experimental items. The d' score reports the subject's ability to discern true old items from the background of new items, effectively accounting for individual response bias. Familiarity was assessed using a composite calculation described by Yonelinas and colleagues (Yonelinas et al. 2010). Familiarity was calculated using the formula: (Know Hit Rate–Know False Alarm Rate)/(1–Recollection), where Recollection = Remember Hit Rate–Remember False Alarm Rate.

Electrodermal activity was analyzed using the same AcqKnowledge software that was used to acquire the data. Preprocessing involved applying a 0.05 Hz high pass filter. Peaks in electrodermal activity following a stimulus were considered responses within the following parameters: a baseline estimation window width of 1 s, threshold for responses was 0.03 µS, and responses under 10% of the maximum were rejected. The minimum separation between the stimulus and event was 0.3 s, and the maximum

separation was 4 s. Once this EDA processing was completed, the average (mean) amplitude, latency, and percentage of stimuli eliciting responses were calculated for each of the three word-types (Pain Mixed, No Pain Mixed, and No Pain Alone) and analyzed using paired samples *t* tests at the 0.05 significance level.

Results

Subject demographics and pain ratings

From the screened subjects meeting the entry criteria, 31 subjects went on to participate in the full experiment. One subject was excluded due to a technical error. Data presented are from a cohort of 30 subjects (22 female) with age in years 23.3 ± 2.8 (mean ± standard deviation). Nerve stimulator intensity in mA was 9.9 ± 3.3. Actual pain ratings obtained just after the practice session were 6.5 ± 0.7 out of 10. Pain ratings following the Pain condition were 6.2 ± 1.1 out of 10.

Psychometric testing

Psychometric testing data were used in exploratory fashion to predict memory performance results, as in previous work (Forkmann et al. 2013, 2016). Plots of the STICA, PVAQ, PASS, and PCS against memory performance for each subject are shown in Supplementary Figs. 3 and 4. Equations describing a line fit through these data are listed in Supplementary Table 2. Surprisingly, scores from these healthy pain-free young adults spanned the scale for all four psychometric tests, indicating a wide variety of preexisting experience with pain and anxiety. The considerable spread in psychometric data in this relatively small cohort of subjects precludes determining any significant relationships. Thus, no differences between experimental conditions based on prior measures of anxiety and pain attitudes can be reliably concluded with our data, and these are presented as exploratory results only.

Response time

Histograms showing RT values before and after transformation are shown in Supplementary Figs. 5 and 6. Response time data are presented with outliers removed (technique described in the Methods). For the RT data, outliers represented 3.7% of learning portion responses and 4.3% of RKN responses.

Response times for the three pain-pairing conditions across the three repetitions in the learning portion are shown in Fig. 2. The RTs (in ms) collapsed across all three word-types during occurrence 1 (mean = 1472, 95%
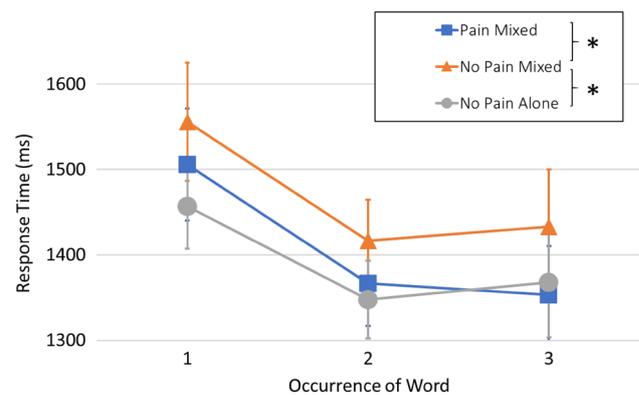

Fig. 2 Response times over the three learning repetitions for each word-type. Significant differences between word-types are indicated with an asterisk (*); this was an overall difference across all three occurrences. Error bars represent standard error

CI = 1374–1581) were significantly greater than occurrence 2 ($P < 0.001$, mean = 1355, 95% CI = 1261–1455) as well as occurrence 3 ($P < 0.001$, mean = 1345, 95% CI = 1253–1445). Collapsing across all three occurrences, differences between word-type can be determined. No Pain Mixed RTs (mean = 1432, 95% CI = 1327–1538) were significantly lower ($P = 0.002$) than No Pain Alone RTs (mean = 1361, 95% CI = 1267–1462). No Pain Mixed RTs were significantly lower ($P = 0.016$) than Pain Mixed RTs (mean = 1374, 95% CI = 1282–1476). There was no significant interaction effect in the LMM between word-type and occurrence number.

The RKN test RT data are shown in Fig. 3. As expected, Know response times were significantly longer than the Remember response times within each word-type (No Pain Mixed, $P = 0.014$, No Pain Alone, $P < 0.001$, and Pain Mixed, $P = 0.002$). Know response times (mean = 1629, 95% CI = 1517–1750) were also significantly longer than Remember response times (mean = 1462, 95% CI = 1361–1567) when collapsed across all three word-types ($P < 0.001$). Know responses for all three word-types were also significantly longer than New response times ($P < 0.001$). When comparing within Remember and Know responses types, there were no significant differences between the three different word-types.

Memory testing

The d' values calculated for Overall and Remember responses are shown in Fig. 4. Panel A compares memory performance for all three word-types for overall responses (Remember and Know) as well as Remember responses separately. There were no significant d' differences between Pain Mixed (mean = 3.33, 95% CI = 3.05–3.61), No Pain Mixed (mean = 3.33, 95% CI = 3.07–3.60) and No Pain

Fig. 3 Response times for Remember (R), Know (K), and New (N) responses during the RKN testing portion, including only correct responses. The number of significant differences prohibits denoting this graphically. R and K responses were significantly different within each word-type, and K responses for each word-type were all significantly different from N responses. Error bars represent standard error
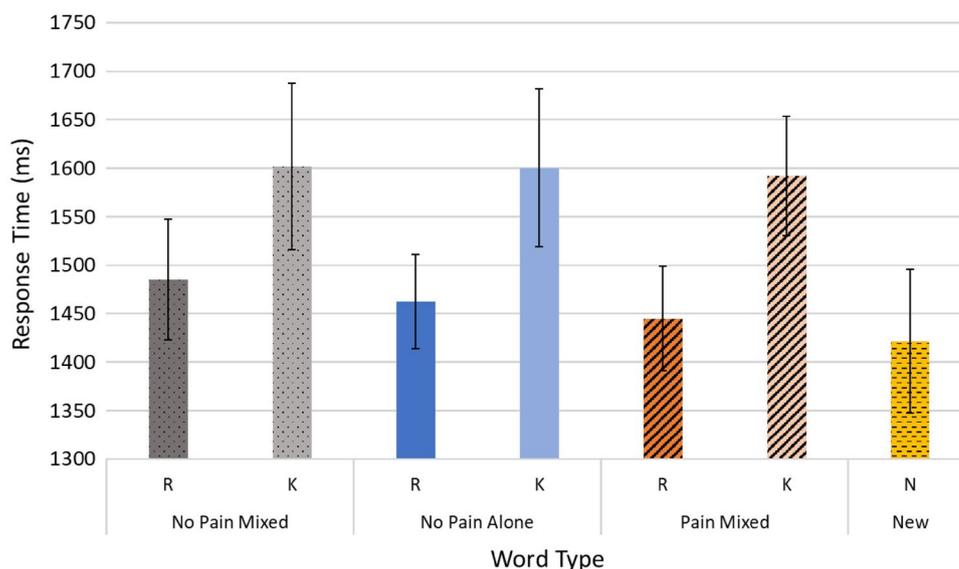
Alone (mean = 3.45, 95% CI = 3.19–3.71) word-type for overall responses. However, for Remember responses, there was a significant effect of pain context, with No Pain Alone words having better recognition than No Pain Mixed words. Specifically, Remember d' was significantly higher ($P = 0.028$) for No Pain Alone words (mean = 2.68, 95% CI = 2.38–2.99) compared to No Pain Mixed words (mean = 2.42, 95% CI = 2.12–2.74). Pain Mixed words (mean = 2.50, 95% CI = 2.17–2.82) d' was not significantly different from either of the other two conditions for Remember responses. Results for familiarity are shown in Fig. 4b. For this composite familiarity measure, there were no significant differences between the Pain Mixed (mean = 0.535, 95% CI = 0.278–0.792), No Pain Mixed (mean = 0.617, 95% CI = 0.546–0.797), and No Pain Alone (mean = 0.483, 95% CI = 0.174–0.792) word-types. For transparency, the hit rate (previously heard words correctly identified) and the false alarm rate (foils incorrectly marked as previously heard) are tabulated in Supplementary Table 1.

The percentage of Remember responses out of total hits is shown in Fig. 5. Although this measure does not account for false-positive responses as does d' or the composite measure of familiarity in Fig. 4, the rates of false-positive responses were not significantly different across experimental word-types. Thus, Fig. 5 graphically shows the proportion of correct responses indicating recollection (Remember) versus familiarity (Know). This proportion was significantly higher ($P = 0.016$) in the No Pain Alone condition (mean = 0.65, 95% CI = 0.56–0.74), compared to the No Pain Mixed condition (mean = 0.57, 95% CI = 0.47–0.67). Pain Mixed was not significantly different than either other condition (mean = 0.60, 95% CI = 0.49–0.70). Notably, when asked at the end of the experiment, no subject could explicitly remember more than four words that were paired with pain,

suggesting that they were not explicitly aware of the association between conditioned and unconditioned stimuli.

When comparing non-pain items between the Pain and No Pain word-types, pain context did affect memory, as stronger recollection is seen for No Pain Alone words, compared to No Pain Mixed words. While the difference in d' for overall responses was not statistically significant between word-types, there were reliably more recollection (Remember) judgments for No Pain Alone words (Fig. 4a). With recollection accounted for, the composite measure of familiarity (Fig. 4b) suggests a corresponding shift to more familiarity judgements for No Pain Mixed words, however, these differences were not statistically significant due to the large variability in familiarity seen in the two other word-types. Figure 5 illustrates that the percentage of total hits with a Remember response is significantly lower for No Pain Mixed words, compared to No Pain Alone words. Pain Mixed words demonstrated intermediate average d' values for Remember responses, but differences between the Pain Mixed words and the two types of non-pain words were not statistically significant. Bearing in mind the potential for a tradeoff between speed versus accuracy during the RKN testing, it is important to note that the RKN response times (Fig. 3) are not demonstrably different across word-types. This suggests that differences in d' are directly reflective of the strength of memory encoding. Taken together, the RKN results indicate that recollection was specifically impaired for the non-pain words that preceded and followed the pain-paired words.

Electrodermal activity

EDA results were tabulated separately for the Pain Mixed and No Pain Mixed word-types, but there were very few

Fig. 4 Recognition memory performance across word-type. **a** Compares average *d'* values for overall recognition (combined Remember and Know responses) and specifically for recollection (Remember responses only). **b** Compares a composite measure of familiarity (see text for details). Error bars display standard error. Significant differences are indicated with an asterisk (*)
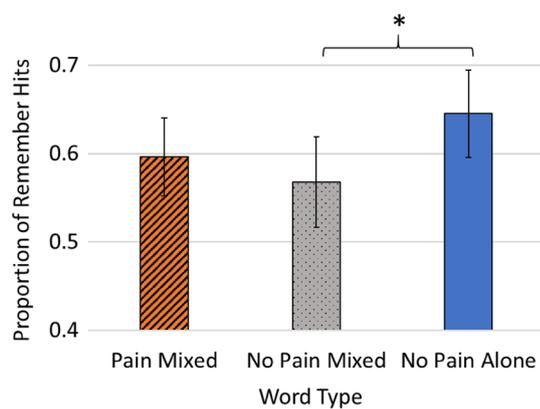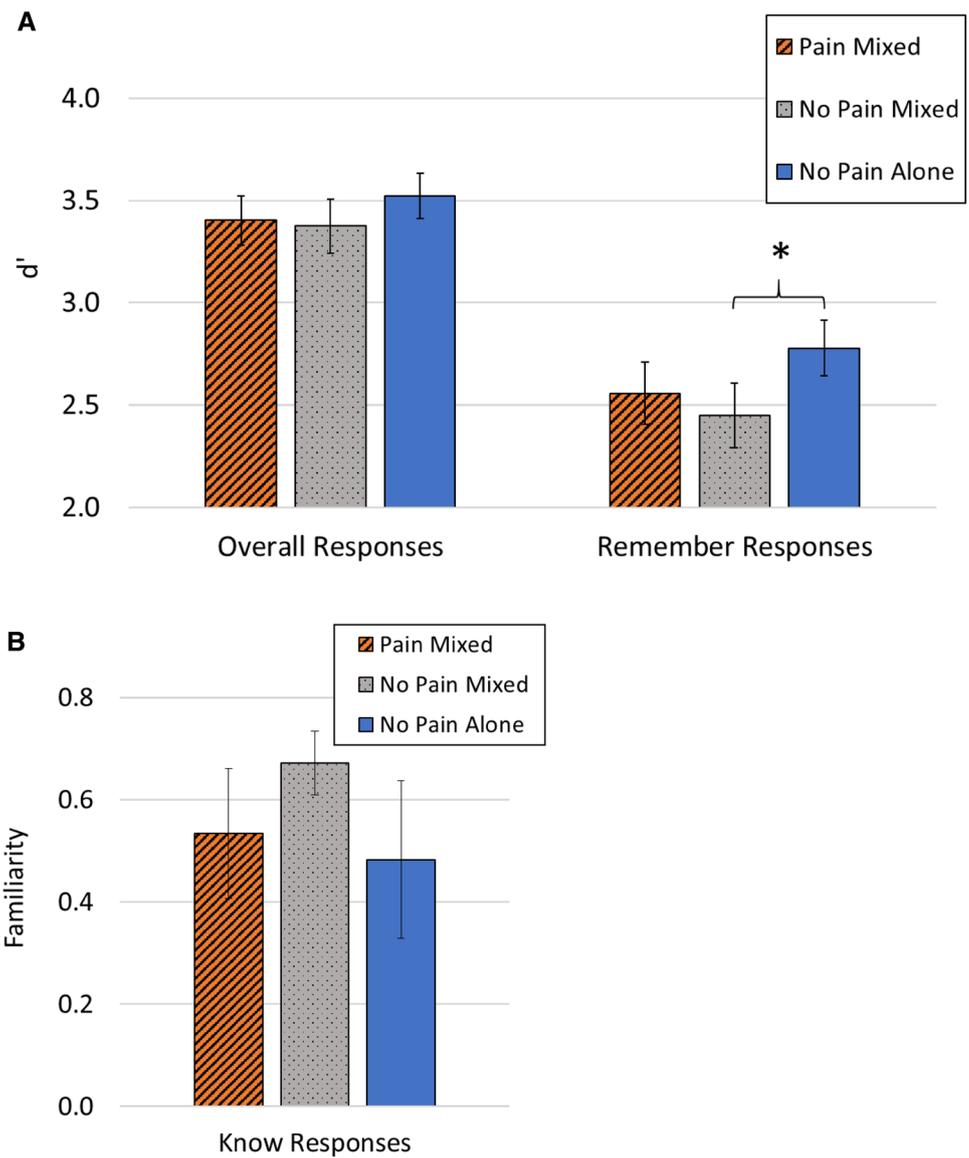
Fig. 5 Proportion of Remember hits out of total hits for each word-type. Error bars display standard error. Significant differences are indicated with an asterisk (*)

EDA responses for the No Pain words. Therefore, only the results from the Pain Mixed word type were analyzed. Most notably, during the RKN testing, less than 10% of all words elicited an EDA response, and the percentage of words with EDA responses did not significantly differ between conditions ($P > 0.05$ for all comparisons). Results for the EDA responses seen during the learning portion of the experiment are presented in the Supplementary Materials.

Discussion

This study primarily examined the effect of acute pain on recognition memory, comparing between pain-paired and non-paired items in two experimental contexts: one with alternating pain stimulations and one free of pain. Counter

to our hypotheses, we found that recollection was higher for non-pain items experienced in the pain-free experimental context (No Pain Alone items), compared to No Pain Mixed items, which were interleaved with pain stimulations. This indicates weaker encoding of non-pain items when preceded and followed by pain-paired items. There were no significant differences in recognition memory between pain and non-pain items within the Pain Condition. Response times for non-pain items within the Pain Condition were slower than other item types. This overall indicates that the effect on memory and task performance from painful stimulations is not limited to the individual pain-paired items, but also affects adjacent non-pain experimental items. We also sought to detect implicit measures of pain memory formation, but were unable to demonstrate significant priming or conditioning effects in the experimental framework employed.

Explicit memory effects

We originally hypothesized that, within the Pain Condition, non-pain items would show stronger recollection than the interleaved pain items. We ultimately found no differences between Pain Mixed and No Pain Mixed words in the Pain Mixed and No Pain Mixed conditions. That is, there were no significant differences in recognition between pain-paired and non-pain items presented in alternating fashion within the same portion of the experiment. Comparing non-pain items between the two conditions, we found better recollection for the No Pain Alone words. Both these findings may be explained by the alternating trial design with respect to shocks in the Pain Condition of the experiment. Working memory resources, critical to bind items into long-term memory, can be consumed by the prior experimental events, affecting performance on the subsequent item (Popov and Reder 2018). Thus, the shock associated with the preceding Pain Mixed word consumed working memory resources, thereby distracting the subject from encoding the subsequent No Pain Mixed word. In effect, this interference likely blunted the recognition memory decrement previously shown for pain-paired items (Forkmann et al. 2016). This aligns with the premise that pain should be arousing; enhancements in memory are not seen when the arousing stimulus is presented in direct attentional competition with the task-relevant experimental items (Mather and Sutherland 2011), as is the paradigm in the current experiment.

The anticipation of an upcoming noxious stimulus has been shown to affect the impact that pain has on performance for adjacent cognitive tasks that are not directly associated with pain. Unpredictability of an experimental pain stimulus is associated with higher ratings of anxiety, negative valence, and intensity (Carlsson et al. 2006). Previous studies have suggested that a warning cue preceding painful

shocks had no effect on explicit memory performance (Forkmann et al. 2016). However, in a study in which subjects received a cue indicating the frequency of shock would be 20, 50, or 80%, a parabolic relationship was demonstrated, with highest memory for items in the 50% probability condition (Bauch et al. 2014). In the present study, the schedule for pain-pairing was consistent and described to subjects in advance, such that the shocks associated with pain-paired words were completely predictable. However, it is possible that the anticipation of a shock on the next item had a distracting effect on the preceding word. Perhaps more likely according to Reder and Popov (Popov and Reder 2018), the previous shock affects encoding of the subsequent word that was not shocked. During the Pain Condition, including both pain-paired and non-pain items, anxiety about the pain-paired word (which is sure to come next) may have worsened encoding for the No Pain Mixed words. It is worth noting that the longest time interval in which no shock will occur starts immediately after a shock is received, including the latter portion of the decision-making period before responding to a pain-paired word. This element of the experimental design may explain some of the unexpected explicit memory results for Pain Mixed words. A revised experiment with pseudorandom timing of pain stimuli is currently underway to further explore the effect for pain-paired experimental events on memory of proximate non-pain items.

Previous studies examining the effect of acute pain on long-term memory have demonstrated different effects, likely due to other variations in experimental design. One factor that may influence performance is subject expectation for the effect of pain on memory. Subjects' preconception that pain would impair memory (rated on a numerical scale) was correlated to decreases in both recollection and familiarity (Forkmann et al. 2016). Second, timing of the experimental pain stimulus relative to incidentally encoded cues may have an impact on the results. Studies in which the painful stimulus was coincident with the experimental stimulus have shown an interruptive effect for pain on memory (Forkmann et al. 2013). Conversely, pain enhanced memory when it followed neutral visual scenes by short delay of 50 ms (Schwarze et al. 2012).

Timing of memory testing may also affect experimental results. Memory enhancement due to pain was only seen in a cohort of subjects with 24-h follow-up testing, compared to a group tested a few minutes after encoding (Schwarze et al. 2012). This suggests that there could be an effect of memory enhancement when allowing time for memory consolidation. However, a second cohort had next-day memory testing, and, except for being in an MRI scanner, they performed the same experiment but showed no difference in memory performance (Schwarze et al. 2012). This raises the possibility of fragility of the results as an explanation of pain's variable effect on long-term memory. This represents

another difference between our study and those demonstrating a pain-related enhancement effect on memory (Schwarze et al. 2012) is the difference in latency period length prior to memory testing. While previous studies had a latency period of 1 day, explicit memory testing in the present study occurred a few minutes after the final part of the encoding experiment. This timeframe and the total number of experimental items being tested (100) exceed the capacity for working memory, such that we are exclusively interrogating long-term memory. However, our design does not allow for any potential consolidation processes to occur. We have another study currently underway to examine the effect that a relatively longer follow-up period would have on memory, as longer time for consolidation in a similar paradigm may show a different result.

Implicit memory effects

We anticipated a decrease in RT to judge a word with each repetition, comparing the first to the subsequent two occurrences (even though the judgment was different on each repetition). This significant decrease over time, demonstrated in Fig. 2, is due in part to the effect of practice on the categorization tasks generally, as well as repetition priming on the lexical items (easier to encode when more familiar). The lack of speed-up from the second and third occurrences suggests that both the task practice effect and the benefit of repeated priming reach a plateau quickly. Of more interest is our demonstration of a significant difference in response times for words in two conditions: the one with pain paired with every other item and the condition with no pain for any items. Words judged on lists that included pain had slower responses than the No Pain Alone condition. This demonstrates an interruptive effect for pain, which seemed to slow subject responses to interleaved non-pain items (No Pain Mixed items).

It is worth pointing out that subjects could have intentionally spaced out the occurrence of shocks by delaying their responses to either or both Pain and No Pain Mixed words. Though shocks immediately followed the pain-paired items, the interval between items (and thus between the one shock and the next) was extended by the subject's response time if it exceeded 1.5 s. Thus, intentional or subconscious delays in responding could have affected the Learning RT results and potentially interfered with memory encoding by cognitive distraction. However, the relatively short response intervals that were obtained (see Fig. 2 and, for more detail, Supplementary Fig. 5) argue against this being an issue. Subjects were also debriefed generally about their experience in the experiment, and none voluntarily reported employing such a strategy (though they were not specifically asked to list strategies employed).

As another measure of implicit learning, we examined conditioned responses to stimuli that had been paired with pain. Though we observed electrodermal responses to the painful stimulations themselves, no conditioned responses were reliably observed during the testing portion of the experiment. One possible explanation is the variation in our experimental design, employing 25 pain-paired (CS+) stimuli, that are repeated 3 times. A more typical conditioning experiment would employ fewer CS+ items, a variable schedule of reinforcement (including some CS- items of same stimulus type), and more repetitions (LeDoux 2014; Lonsdorf et al. 2017). Though there certainly are experimental paradigms in which both classical conditioning and explicit memory responses can be elicited (Dunsmoor and Kroes 2019), these two effects are probably not best assessed concomitantly with an experimental design like the present study.

Limitations

We believe that recognition follows a dual-process model, with recollection and familiarity representing distinct aspects of explicit memory (Diana et al. 2006). The use of the Remember-Know paradigm to probe recollection versus familiarity is not without criticism (Wais et al. 2008). However, much of the challenge centers on how the RKN task instructions are given to the participants (Migo et al. 2012). In the current experiment, we have taken great care to ensure consistent instructions were given to the subjects, describing Remember responses as those with explicit recollection and Know responses as those with familiarity only. This included reviewing the written instructions (a copy of which is included in the Appendix) with them, as well as reiterating the differences with pre-recorded instructions just before the task commenced. Also worthy of note is that Remember and Know judgments have similar ROC curves (Arndt and Reder 2002) and are underpinned by distinct cerebral resources (Migo et al. 2012). Our composite measure of familiarity, adapted from (Yonelinas et al. 2010), should eliminate the bias in Know responses to items for which recollection is not achieved.

Putting our explicit memory results in the context of the broader literature, one could interpret our equivocal differences specifically for pain-paired items, compared to both types of non-pain items, as a null result. However, the stronger recollection performance for No Pain Alone versus No Pain Mixed words (neither of which are immediately associated with painful stimulation) demonstrates that pain context affects recognition memory. This suggests that, at least in a 1:2 pain-pairing paradigm, residual effects from pain-paired items reduce the strength of memory encoding for the interleaved No Pain Mixed items.

Conclusions

For aurally delivered words, about which subjects performed a categorization decision, experimental painful stimulation affected memory performance. Explicit memory, indicated by Remember-Know testing, was most affected. Compared to non-pain words heard in a pain-free portion of the experiment, non-pain words heard in the context of other pain-associated words were not as deeply encoded. This was demonstrated by less recollection. An important accompanying finding is that response times to repeated experimental items were slower for non-pain words heard in the context of pain-paired words, compared to non-pain items heard in a pain-free context. Taken together, these results suggest that, compared to a pain-free period, acute pain stimulations create an experiential context wherein memory formation is affected, even for items not directly paired with painful stimulations.

References

Andreatta M, Leombruni E, Glotzbach-Schoon E, Pauli P, Muhlberger A (2015) Generalization of contextual fear in humans. Behav Ther 46:583–596. https://doi.org/10.1016/j.beth.2014.12.008

Arndt J, Reder LM (2002) Word frequency and receiver operating characteristic curves in recognition memory: evidence for a dual-process interpretation. J Exp Psychol Learn Mem Cogn 28:830–842

Bauch EM, Rausch VH, Bunzeck N (2014) Pain anticipation recruits the mesolimbic system and differentially modulates subsequent recognition memory. Hum Brain Mapp 35:4594–4606. https://doi.org/10.1002/hbm.22497

Bennion KA, Ford JH, Murray BD, Kensinger EA (2013) Oversimplification in the study of emotional memory. J Int Neuropsychol Soc 19:953–961. https://doi.org/10.1017/S1355617713000945

Berryman C, Stanton TR, Jane Bowering K, Tabor A, McFarlane A, Lorimer Moseley G (2013) Evidence for working memory deficits in chronic pain: a systematic review and meta-analysis. Pain 154:1181–1196. https://doi.org/10.1016/j.pain.2013.03.002

Bingel U, Quante M, Knab R, Bromm B, Weiller C, Buchel C (2002) Subcortical structures involved in pain processing: evidence from single-trial fMRI. Pain 99:313–321

Carlsson K, Andersson J, Petrovic P, Petersson KM, Ohman A, Ingvar M (2006) Predictability modulates the affective and sensory-discriminative neural processing of pain. Neuroimage 32:1804–1814

Castegnetti G, Tzovara A, Staib M, Paulus PC, Hofer N, Bach DR (2016) Modeling fear-conditioned bradycardia in humans. Psychophysiology 53:930–939. https://doi.org/10.1111/psyp.12637

Cave CB, Squire LR (1992) Intact and long-lasting repetition priming in amnesia. J Exp Psychol Learn Mem Cogn 18:509–520

Choi JC, Park SK, Kim YH et al (2006) Different brain activation patterns to pain and pain-related unpleasantness during the menstrual cycle. Anesthesiology 105:120–127

D'Esposito M (2007) From cognitive to neural models of working memory. Philos Trans R Soc Lond B Biol Sci 362:761–772. https://doi.org/10.1098/rstb.2007.2086

Diana RA, Reder LM, Arndt J, Park H (2006) Models of recognition: a review of arguments in favor of a dual-process account. Psychon Bull Rev 13:1–21

Dunsmoor JE, Kroes MC (2019) Episodic memory and Pavlovian conditioning: ships passing in the night. Curr Opin Behav Sci 26:32–39. https://doi.org/10.1016/j.cobeha.2018.09.019

Forkmann K, Wiech K, Ritter C, Sommer T, Rose M, Bingel U (2013) Pain-specific modulation of hippocampal activity and functional connectivity during visual encoding. J Neurosci 33:2571–2581

Forkmann K, Schmidt K, Schultz H, Sommer T, Bingel U (2016) Experimental pain impairs recognition memory irrespective of pain predictability. Eur J Pain 20:977–988. https://doi.org/10.1002/ejp.822

Francis WS (2014) Repetition priming in picture naming: sustained learning through the speeding of multiple processes. Psychon Bull Rev 21:1301–1308. https://doi.org/10.3758/s13423-014-0610-9

Grillon C (2002) Associative learning deficits increase symptoms of anxiety in humans. Biol Psychiatry 51:851–858

Gros DF, Antony MM, Simms LJ, McCabe RE (2007) Psychometric properties of the State-Trait Inventory for Cognitive and Somatic Anxiety (STICSA): comparison to the State-Trait Anxiety Inventory (STAI). Psychol Assess 19:369–381. https://doi.org/10.1037/1040-3590.19.4.369

Henke K (2010) A model for memory systems based on processing modes rather than consciousness. Nat Rev Neurosci 11:523–532. https://doi.org/10.1038/nrn2850

Hood A, Pulvers K, Spady TJ (2013) Timing and gender determine if acute pain impairs working memory performance. J Pain 14:1320–1329. https://doi.org/10.1016/j.jpain.2013.05.015

Keane MM, Cruz ME, Verfaellie M (2015) Attention and implicit memory: priming-induced benefits and costs have distinct attentional requirements. Mem Cognit 43:216–225. https://doi.org/10.3758/s13421-014-0464-4

LaBar KS, Cabeza R (2006) Cognitive neuroscience of emotional memory. Nat Rev Neurosci 7:54–64. https://doi.org/10.1038/nrn1825

LeDoux JE (2014) Coming to terms with fear. Proc Natl Acad Sci USA 111:2871–2878. https://doi.org/10.1073/pnas.1400335111

Leys C, Ley C, Klein O, Bernard P, Licata L (2013) Detecting outliers: do not use standard deviation around the mean, use absolute deviation around the median. J Exp Soc Psychol 49:764–766. https://doi.org/10.1016/j.jesp.2013.03.013

Lonsdorf TB, Menz MM, Andreatta M et al (2017) Don't fear 'fear conditioning': methodological considerations for the design and analysis of studies on human fear acquisition, extinction, and return of fear. Neurosci Biobehav Rev 77:247–285. https://doi.org/10.1016/j.neubiorev.2017.02.026

Lund K, Burgess C (1996) Producing high-dimensional semantic spaces from lexical co-occurrence. Behav Res Methods, Instrum Comput 28:203–208. https://doi.org/10.3758/bf03204766

Mather M, Sutherland MR (2011) Arousal-biased competition in perception and memory. Perspect Psychol Sci 6:114–133. https://doi.org/10.1177/1745691611400234

McCracken LM, Dhingra L (2002) A short version of the Pain Anxiety Symptoms Scale (PASS-20): preliminary development and validity. Pain Res Manag 7:45–50

Migo EM, Mayes AR, Montaldi D (2012) Measuring recollection and familiarity: improving the remember/know procedure. Conscious Cogn 21:1435–1455. https://doi.org/10.1016/j.concog.2012.04.014

Murty VP, Adcock AR (2017) Distinct medial temporal lobe network states as neural contexts for motivated memory formation. In: Hannula DE, Duff MC (eds) The hippocampus from cells to systems: structure, connectivity, and functional contributions to memory and flexible cognition. Springer International Publishing, Cham, pp 467–501

Osman A, Barrios FX, Kopper BA, Hauptmann W, Jones J, O'Neill E (1997) Factor structure, reliability, and validity of the pain catastrophizing scale. J Behav Med 20:589–605

Popov V, Reder LM (2018) Frequency effects on memory: a resource-limited theory. https://osf.io/dsx6y/

Rainville P, Duncan GH, Price DD, Carrier B, Bushnell MC (1997) Pain affect encoded in human anterior cingulate but not somatosensory cortex. Science 277:968–971

Rajaram S (1993) Remembering and knowing: two means of access to the personal past. Mem Cognit 21:89–102

Roelofs J, Peters ML, McCracken L, Vlaeyen JW (2003) The pain vigilance and awareness questionnaire (PVAQ): further psychometric evaluation in fibromyalgia and other chronic pain syndromes. Pain 101:299–306

Schiller D, Monfils MH, Raio CM, Johnson DC, Ledoux JE, Phelps EA (2010) Preventing the return of fear in humans using reconsolidation update mechanisms. Nature 463:49–53. https://doi.org/10.1038/nature08637

Schreckenberger M, Siessmeier T, Viertmann A et al (2005) The unpleasantness of tonic pain is encoded by the insular cortex. Neurology 64:1175–1183

Schwarze U, Bingel U, Sommer T (2012) Event-related nociceptive arousal enhances memory consolidation for neutral scenes. J Neurosci 32:1481–1487. https://doi.org/10.1523/JNEUROSCI.4497-11.2012

Shields GS, Sazma MA, McCullough AM, Yonelinas AP (2017) The effects of acute stress on episodic memory: a meta-analysis and integrative review. Psychol Bull 143:636–675. https://doi.org/10.1037/bul0000100

Squire LR, Dede AJ (2015) Conscious and unconscious memory systems. Cold Spring Harb Perspect Biol 7:a021667. https://doi.org/10.1101/cshperspect.a021667

Wais PE (2008) FMRI signals associated with memory strength in the medial temporal lobes: a meta-analysis. Neuropsychologia 46:3185–3196. https://doi.org/10.1016/j.neuropsychologia.2008.08.025

Wais PE, Mickes L, Wixted JT (2008) Remember/know judgments probe degrees of recollection. J Cogn Neurosci 20:400–405. https://doi.org/10.1162/jocn.2008.20041

Yonelinas AP, Ritchey M (2015) The slow forgetting of emotional episodic memories: an emotional binding account. Trends Cogn Sci 19:259–267. https://doi.org/10.1016/j.tics.2015.02.009

Yonelinas AP, Aly M, Wang WC, Koen JD (2010) Recollection and familiarity: examining controversial assumptions and new directions. Hippocampus 20:1178–1194. https://doi.org/10.1002/hipo.20864

Zanto TP, Clapp WC, Rubens MT, Karlsson J, Gazzaley A (2016) Expectations of task demands dissociate working memory and long-term memory systems. Cereb Cortex 26:1176–1186. https://doi.org/10.1093/cercor/bhu307