Involuntary Memories of Emotional Scenes: The Effects of Cue Discriminability and Emotion over Time

Søren R. Staugaard & Dorthe Berntsen

Center on Autobiographical Memory Research, Department of Psychology and Behavioral Sciences, Aarhus University

Corresponding author:

Søren Risløv Staugaard

Center on Autobiographical Memory Research

Department of Psychology and Behavioral Sciences

Aarhus University

Bartholins Allé 9

8000 Aarhus C. Denmark

Email: sorsta@psy.au.dk

Abstract

Involuntary episodic memories come to mind spontaneously – that is, with no preceding retrieval attempts. Such memories are frequent in daily life, in which they are predominantly positive and often triggered by situational features matching distinctive parts of the memory. However, individuals suffering from psychological disorders, such as Posttraumatic Stress Disorder (PTSD), have stressful, repetitive and unwanted involuntary memories about negative events in their past. These unwanted recollections are disturbing and debilitating. Although such intrusive involuntary memories are observed across a range of clinical disorders, there is no broadly agreed upon explanation of their underlying mechanisms, and no successful experimental simulations of their retrieval. In a series of experiments, we experimentally manipulated the activation of involuntary episodic memories for emotional and neutral scenes and predicted their activation on the basis of manipulations carried out at encoding and retrieval. Our findings suggest that the interplay between cue discriminability at the time of retrieval and emotional arousal at the time of encoding are crucial for explaining intrusive memories following negative events. While cue distinctiveness is important directly following encoding of the scenes, emotional intensity influences retrieval after delays of 24 hours and a week. Voluntary remembering follows the same pattern as involuntary remembering. Our results suggest an explanatory model of intrusive involuntary memory for emotional events in clinical disorders.

Involuntary Memories of Emotional scenes: The Effects of Cue Discriminability and Emotion over Time

A male survivor of an assault recounts an evening where he is lying in his bed surrounded by darkness and trying to go to sleep. He is relaxed and at peace, when he notices the light being turned on in the neighboring building. This is followed by an involuntary memory of the painful moment when he regained consciousness after the assault: "I wake up because light is shining into my eyes and I hear voices around me. A stranger helps me to sit up. I am confused and my head is hurting. When I touch it, my hand gets filled with blood. When I look around, I also see blood on the floor and I then understand why my head is hurting. People help me out of the dark room where I have been lying till this moment. There is a lot of fuss around me, but I just look at my hands, not understanding what has happened." (from Berntsen, 2001).

Survivors of trauma often experience vivid, intrusive memories of the traumatic event. Such intrusive memories are considered a hallmark of Posttraumatic Stress Disorder (PTSD). They were classified as one of the defining criteria for PTSD when the diagnosis was introduced in the DSM-III and have been maintained in all subsequent revisions of the diagnostic manual (American Psychiatric Association, 1980, 1987, 1994, 2000, 2013). Intrusive images of stressful autobiographical events are also found in many other disorders, such as depression, eating disorders, anxiety disorders, and bipolar disorder (for reviews see Brewin, Gregory, Lipton, & Burgess, 2010; Steel & Holmes, 2007). In spite of their presence across a range of disorders, well-established explanations of the underlying mechanisms are lacking. One likely reason for this shortcoming is their spontaneous and uncontrollable nature, which makes controlled laboratory experimentation difficult.

In this article, we describe a series of experiments intended to simulate real-life intrusive memories in a laboratory setting. Our method presents a major advantage over previous work by controlling both the encoding and retrieval situation, which allows us to predict the occurrence of intrusive involuntary memories by manipulating their cues, their emotional content as well as their retention interval. We present evidence suggesting that the ways in which these factors interact may be of central importance to the development of intrusive memories after negative events.

Mechanisms Underlying Involuntary Autobiographical Memories

Involuntary memories are not limited to traumatic events, but are common in daily life (e.g., Ball & Little, 2006; Berntsen, 1996; Berntsen & Hall, 2004; Kvavilashvili & Mandler, 2004; Mace, 2004; for reviews see Mace, 2007; Berntsen, 2009, 2010). Most everyday involuntary memories are emotionally positive (for an overview see Berntsen, 2009). An explanatory model for the activation of everyday involuntary autobiographical memories was proposed by Berntsen (2009) and elaborated by Berntsen, Staugaard, and Sørensen (2013). This model integrates findings from the autobiographical memory literature with well-established findings of cued recall from standard laboratory memory research (for a review, see Roediger & Guynn, 1996), notably the principle of encoding specificity (Tulving & Thompson, 1973) and the principle of cue overload (Watkins and Watkins, 1975).

According to the principle of encoding specificity, the informational overlap between a cue and an item to be remembered predicts the likelihood that the item is retrieved in response to the cue. Although this principle was introduced to explain cued recall in standard word list experiments, several theories have adhered to this notion when trying to explain the spontaneous activation of autobiographical memories in naturalistic settings (e.g., Conway, 2005; Moscovitch, 1995). This is because the great majority of involuntary autobiographical memories seem to be facilitated by cues in the retrieval context (Berntsen, 2009). However, when applied to everyday involuntary memories, the notion of encoding specificity appears to have difficulties explaining why we are not constantly flooded by involuntary memories of past events: Most situations that we encounter throughout the day by necessity include a great number of overlapping features with events we have encountered some time in the past, and we are able to voluntarily generate memories in response to many of such environmental features, if instructed to do so (Galton, 1907; Berntsen & Hall, 2004). Thus, if involuntary autobiographical memories were simply contingent upon an encoding-retrieval match, it seems that we should be flooded by such memories throughout our waking life (Berntsen, 2009).

There are at least two possible factors, which may contribute to this effect. One is that past events are not equally salient in memory. As pointed out by Tulving (1974), memory for a given event is a product of both the present cue and the memory trace itself, as this has been laid down as a result of the original experience of the event. The more emotional, uncommon, recent or otherwise salient the event is in memory, the more likely it is to be evoked by a given cue. This entails that not all memories have an equal likelihood of entering consciousness in response to a matching cue. When hearing the siren of an approaching ambulance, you are unlikely to remember just any occasion when you heard an ambulance, but highly likely to remember the time you were in the back of one.

The second factor relates to the cueing of the memories and involves supplementing the notion of encoding specificity with the notion of cue overload stating that “The probability of recalling an item declines with the number of items subsumed by its functional retrieval cue” (Watkins & Watkins, 1975, p. 442). Thus, the likelihood of a cue providing access to a given target memory depends on the extent to which this cue is uniquely associated with the target. Its strength declines to the extent it is associated with other memories as well. Applied to an everyday memory setting this would mean that if the overlapping feature between a present and a past event is a rare one (e.g., a unique taste, or a specific piece of music encountered only once before), the probability of spontaneously recalling the event is enhanced. Integrating these principles in relation to memory for counting-out rhymes and ballads, Rubin (1995) introduced the notion of cue-item discriminability. He defined it as “how easily a given cue isolates an item. … Simply put, a word is likely to be recalled if, on the basis of the cues available at the time, it can be discriminated from all else in memory” (p. 146).

On the basis of these well-established principles from laboratory memory research in combination with findings from studies on involuntary autobiographical memories in everyday life, Berntsen et al. (2013) developed an experimental paradigm in which the activation of involuntary memories was controlled and predicted on the basis of manipulations done at encoding. Using a paired-associate paradigm, the participants were presented with pictures of mundane scenes paired with sounds. Both scene and sound could be either unique (derived from a category that was presented only once) or repeated (derived from a category that was presented several times). During retrieval, the participants conducted an attention demanding sound location task employing sounds from the encoding phase. In addition to the sound location task, they were asked to record all memories that might spontaneously arise during this task. Consistent with the notion of cue overload, unique sounds coupled with unique pictures generated most involuntary memories. However, in this series of studies all scenes were emotionally neutral in order to examine the effect of the distinctiveness of the cues. The important question here is whether highly emotional scenes will follow the same basic mechanism or if the emotional intensity will assign priority to these scenes in involuntary retrieval even when the associated cues are repeated and therefore less distinctive? In other words, we want to examine the interplay between the emotionality of the event and the distinctiveness of the associated cues in order to arrive at a deeper understanding of involuntary, intrusive memories of stressful or traumatic events.

We include a condition of voluntary (strategic) recall in response to the same sample of cues, in order to be consistent with previous work (Berntsen et al., 2013) and in order to provide a comparison with the involuntary condition. We expect the voluntary condition to involve more retrieval effort but to show the same effects of emotion at encoding as the involuntary memories. This would be consistent with the view that both types of retrieval operate on the same underlying memory structures (e.g., Berntsen, 2009, 2010).

Intrusive Involuntary Memories of Emotional Events

Clinical anecdotes suggest that highly emotional events may indeed be triggered by repeated and mundane cues. There are numerous case reports of patients experiencing stressful flashbacks of traumatic events in response to cues such as utensils (Salomons, Osterman, Gagliese, & Katz, 2004), the smells of diesel fuel and salt water (Sierra & Berrios, 1999), a bright patch of sunlight on a lawn (Ehlers, Hackmann, & Michael, 2004), and the smell of a barbecue (Berntsen et al., 2013; see also the example in the beginning of this article). A woman employed in a hospital reported intense flashbacks when confronted with blue scrub suits, and eventually the colour blue itself (Salomons et al., 2004). If the distinctiveness of the cue is decisive, such common features should not be very efficient at triggering involuntary memories, yet in the anecdotes they appear to be.

The key to explaining this phenomenon may be the enhanced salience and persistency of emotional events in memory. A wealth of research has demonstrated that emotional events are persistent in memory (e.g., Cahill & McGaugh, 1998; McGaugh, 2003; Porter & Peace, 2007). In addition, emotional memories are retrieved with more detail compared with non-emotional memories (Schaefer & Philippot, 2005; St. Jacques & Levine, 2007) and with a greater sense of vividness and recollection (Talarico, LaBar, & Rubin, 2004). Several theoretical accounts offer explanations for why emotional events are superior to non-emotional events in memory. The interference theory of forgetting states that learning of new events requires the inhibition of memories of events - acquired either prior to encoding or during the retention interval - that share some similarity with the new events (Underwood, 1957; Underwood & Postman, 1960, see Bjork, 2003, for a recent review). Over time, inhibited memories can recover, causing interference with memory for the new event. Highly emotional events, however, are less likely to show similarity with prior or newly acquired events and are therefore less likely to show forgetting due to interference. Some theorists have argued for an updated account of interference focusing more on consolidation (e.g. Wixted, 2004). Within this theory, the interfering events do not have to be similar to the memory, since establishing new memories disrupts consolidation of recent memories by itself (for a review, see Wixted, 2004).

At a neurophysiological level, consolidation has been researched in different animals (McGaugh, 2004). In this model, emotionally arousing events increase the flow of noradrenaline and cortisol to the basolateral nucleus of the amygdala. This process initiates synaptic changes in the hippocampus and other brain areas involved in memory, eventually altering the expression of the neurons' genes. This consolidation of the emotional memory can be divided into short and long phases, corresponding to short- and long-term memory (Dudai, 2004). The short phase consolidates the memory within minutes to hours after the event, while the long phase requires 24 hours or more. Given the stronger consolidation of emotional memories compared with non-emotional memories, emotional events should be more resistant to interference over time.

Mirroring these neural processes, cognitive and socio-cultural theories have provided explanations for the increased accessibility of emotional memories focusing on rehearsal and greater centrality to life story and identity (Berntsen & Rubin, 2008). Irrespective of their level of analysis, the theories share the central prediction that highly emotional events may retain increased salience in memory. Therefore they may be brought to mind even when a cue is weak, that is, when the cue is associated with many other past events. Returning to the anecdote provided in the beginning of the article, seeing a light being turned on in a dark room was presumably associated with numerous past events. None of these were probably as salient as the experience of regaining consciousness after the assault. This event was likely to be better consolidated, more rehearsed, and more central to the person's life-story and identity. For all of these reasons, this particular episode was likely to be triggered by the light, although this cue was not distinctive.

The Present Series of Experiments

In the present experiments we attempted to experimentally manipulate the activation of intrusive memories of emotional scenes. Many previous studies using the trauma analogue paradigm (e.g., Ehlers, Mauchnik, & Handley, 2012; Holmes, Brewin, & Hennessy, 2004; Pearson, Ross, & Webster, 2012) have controlled the encoding context by showing emotional films or pictures. However, in these studies the retrieval phase has typically been examined through a diary technique in natural settings, leaving fewer possibilities for experimental control. Here, we control both the encoding and the retrieval phase through a paired-associate laboratory technique developed in our prior work (Berntsen et al., 2013). We aim to show that the activation of intrusive memories during the retrieval phase can be predicted on the basis of manipulations conducted during encoding, i.e., whether the cue is distinctive or non-distinctive and whether the scene is emotional or non-emotional.

Based on the literature review, we have two predictions. First, the distinctiveness of a cue should lead to a higher frequency of involuntary memories, as we have shown previously (Berntsen et al., 2013). Second, memories of emotional scenes should be more accessible to retrieval compared with neutral events regardless of the distinctiveness of the associated cues. If we assume that these two forces act on retrieval frequency with the same strength, emotional events associated with distinctive cues should have the greatest advantage due to an additive effect of their respective influences, while neutral events associated with non-distinctive cues should be quite infrequent.

Modifying the method used in prior research (Berntsen et al., 2013), we wanted to investigate the interacting effects of cueing and emotionality of events on involuntary memories. Participants first viewed a series of emotional or neutral naturalistic scenes, each paired with a sound that was either unique (i.e., derived from a category of sounds that was only presented once) or repeated (i.e., derived from a category of sounds that was presented several times). Subsequent to this encoding task, the participants conducted a target location task while hearing the sounds from encoding mixed with some novel sounds. While doing the target location task, they were asked to also indicate if any of the previously presented scenes spontaneously came to mind. The participants were not informed that they were participating in a memory experiment (see Berntsen et al., 2013, for a detailed description of the rationale behind this procedure and how it is assumed to map onto the activation of involuntary memories in real life).

We included a voluntary retrieval condition as a comparison. We expected the voluntary condition to be similarly affected by cue distinctiveness and emotion, consistent with the theoretical assumption that both types of retrieval operate on the same underlying memory system (e.g., Berntsen, 2009, 2010; Berntsen & Rubin, 2014). However, we expected voluntary retrieval to be more effective, thus generating higher frequencies of memories across emotion and cueing conditions compared with the involuntary condition (Berntsen et al., 2013). We also expected longer retrieval times for the voluntary memories, in accordance with prior findings (Berntsen et al., 2013) and with the hypothesis that voluntary retrieval requires more executive processing than involuntary retrieval.

The local Committee on Health Research Ethics approved the use of highly negative images within the present paradigm in healthy, consenting adults. We obtained written consents from all participants in the experiments. All participants understood that participation was completely voluntary, that they could resign from the experiment at any time, and that all data were anonymous and confidential. To compensate them for their time, all participants in all experiments were given cinema tickets. As part of the debriefing procedures, participants were offered the opportunity to contact the first author with any questions or concerns related to the study. However, none of the participants made use of this offer in any of the experiments.

Experiment 1

Participants

Thirty-two participants (25 female, mean age 22.31, range 19-38) were recruited from Aarhus University's student population and randomly assigned to either voluntary or involuntary retrieval.

Design

The design was a 2 (Emotion: emotional *vs.* neutral scene) x 2 (Cueing: unique *vs.* repeated sound) x 2 (Retrieval: involuntary *vs.* voluntary condition) mixed design with Emotion and Cueing as within subjects variables and Retrieval as a between-groups variable.

Materials

Thirty-two pictures of scenes were taken from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2008). Sixteen of them were of highly negative emotional valence according to the ratings included with the set (valence: M = 1.81, SD = 0.09; arousal: M = 6.17, SD = 0.42) while the other 16 were neutral (valence: M = 4.57, SD = 0.11; arousal: M = 3.27, SD = 0.48). Examples of the highly negative scenes included severely scarred or mutilated bodies and children starving. The neutral scenes included people in neutral settings as well as objects. In addition, 16 pictures of scenes were taken from various sources on the internet. Each of these pictures was selected to carefully match an individual IAPS picture in terms of visual similarity and content to be used as foils in a recognition task. All pictures were rescaled to 1024 x 768 pixels.

Forty-eight sounds were taken from a previous study (Berntsen et al., 2013) as well as from various royalty-free sound libraries. All sounds were normalized and cut short at four seconds. Alternate versions were created of all sounds by panning them 75 % to the right or the left. While 16 of the sounds were dissimilar (e.g., a person coughing or a telephone ringing), of the remaining 32 sounds, there were four sets of highly similar sounds (birds singing, dogs barking, car engines revving and busy pedestrian streets). All sounds were presented only once during encoding, but the dissimilar sounds would appear unique to the listener while the similar sounds would appear repeated, since they derived from the same category (e.g., dogs barking) and were difficult to distinguish from one another. All sounds were selected to be emotionally neutral and indeed a panel of independent judges (N = 9) rated them as such (M = 2.89, SD = 0.27, on a scale of 1-5, with 1 = highly negative, 3 = neutral and 5 = highly positive). Finally, an additional 32 sounds were collected using the same procedure described above, to be used as filler sounds in the involuntary condition.

Method

A computerized task was programmed in E-Prime 2.0 Professional (Psychology Software Tools, Inc). The task was based on a previous version by Berntsen et al. (2013), but was modified to include emotional scenes. The task consisted of three different phases: 1) An encoding phase, where participants would hear different sounds each paired with a scene; 2) a retrieval phase, where participants would hear the same sounds again and either a) attempt to deliberately recall the scenes (voluntary condition), or b) perform a simple attention task while registering if a memory of a scene spontaneously came to mind (involuntary condition); and 3) a recognition phase, where participants were presented with the scenes from the encoding phase paired with foils, and asked to decide which one they had previously seen. Only the retrieval phase differed between the two conditions. Each phase is described in detail below.

Encoding Phase

During encoding, participants were shown 32 trials of scenes paired with sounds. On each trial, a sound and a scene were presented simultaneously for four seconds. Participants were then asked to indicate on a 5-point scale how emotionally intense they thought the scene was, with 1 being "Not at all intense" to 5 being "Very intense". Next, the same scene was presented alone, and participants were asked to try to remember the sound that it was paired with. They then were asked to rate the relatedness of the sound and the scene in the pair. Specifically, the question was: "To what extend do you think that the sound and the scene could form part of the same story?" and was answered on a 5-point scale from 1 = "Not at all" to 5 = "Very well". The trial then proceeded to the next sound-scene pair.

The encoding phase was divided into four different, unblocked types of trials depending on whether the scene was neutral versus emotional, and whether the sound was unique versus repeated. On one type of trial, a unique sound cue was paired with an emotional scene. As shorthand, this trial is referred to as UniCue_EmoScene. On a second type of trial, a unique sound was paired with a neutral scene (UniCue_NeuScene). The third type of trial was one category of repeated sounds (e.g., birds singing or car engines revving) paired with an emotional scene (RepCue_EmoScene), and the fourth type of trial was another category of repeated sounds (e.g., dogs barking or busy pedestrian streets) paired with a neutral scene (RepCue_NeuScene). In total, participants saw eight trials of each of these four types. The order of presentation was randomized so that the same sound was not paired with the same scene for every participant. Finally, we counterbalanced the pairings between repeated sound category and valence of the scenes.

Retrieval Phase

In the involuntary condition, participants would hear each of the 32 sounds from the encoding phase presented once in each ear (i.e., to the left or to the right). In addition, they would hear 32 unfamiliar sounds presented once to each ear. These sounds were selected to be of similar categories to the familiar sounds (yet easily distinguishable from them), such as animal, human, and natural sounds, and they would also appear unique or repeated to the listener. In total, participants in the involuntary condition completed 128 trials. On each trial, the sound was presented together with a fixation cross on the screen. One and a half seconds into the playback of the sound, a bright star would appear in either the left or the right side of the screen. The participants were asked to indicate in which side of the screen the star was located by pressing “1” for left and “2” for right. This simple attention task was used as a cover task to minimize the risk that participants would deliberately search for memories of scenes, but also to simulate the conditions under which involuntary memories are most likely to appear (e.g., during monotonous or non-demanding tasks, e.g., see Berntsen, 1998, 2009; Schlagman & Kvavilashvili, 2008; Singer, 1966). While performing this cover task, they were told to indicate whenever they spontaneously remembered a scene from the encoding phase. If so, they should immediately press “3”, which would then display a digital questionnaire on the screen. The questionnaire included space to write a brief description of the retrieved scene and three questions to be answered on a 5-point scale. The first question concerned the specificity of the retrieved scene (from 1 = “Not specific at all” to 5 = “Highly specific”); the second question concerned the emotional intensity of the retrieved scene (from 1 = “Not intense at all” to 5 = “Very intense”); and the third question concerned bodily reactions to the retrieved scene, such as increased heart rate, nervousness or tension (from 1 =

"Not at all" to 5 = "Very much"). Sounds were presented in a fixed random order, making sure that a particular type of trial was not repeated more than three times in a row.

In the voluntary condition, participants heard the 32 sounds from the encoding phase once presented to both ears (i.e., centered). They were asked to recall the scene that was paired with the sound during the encoding phase. If they did so, they were to press "1". If they could not remember a scene, they were to press "2". When participants indicated that they had recalled a scene, they were presented with the same questionnaire as in the involuntary condition. The sounds were ordered to ensure that the same type of trial was not repeated more than twice in a row.

The reason for the unequal number of cues in the involuntary versus voluntary condition was the fact that the involuntary memories by definition are activated by chance in an 'accidental' fashion for which reason only some of the cues would be likely to be followed by a memory. The goal-directed search used in the voluntary condition, on the other hand, was more likely to generate memories in response to most of the cues. Having a high number of cues in the voluntary condition would prolong this condition relative to the involuntary condition and potentially reduce the quality of the responses. The current design is similar to the ones used in previous studies (Berntsen et al., 2013; Schlagman & Kvavilashvili, 2008).

Recognition Phase

All participants reviewed the 32 scenes from the encoding phase one at a time. This time, each scene was presented alongside a foil. Participants were asked to indicate which of the two scenes they had seen before, as well as rate the confidence in their own response on a 5-point scale (From 1 = "Very uncertain" to 5 = "Very certain"). The trials were presented in a counterbalanced, fixed order, where the foils would be on the left side of the screen half the time, while also making sure that the same type of trial was not repeated more than twice in a row. The purpose of the recognition task was to check whether participants actually looked at all of the scenes during the encoding phase, since some might feel inclined to “look past” the negative scenes due to their content. Each foil was selected to carefully match the corresponding IAPS scene in terms of content and composition. The difficulty of the recognition task differed from trial to trial with some foils being better matches than others, but in general it was not difficult as long as participants paid attention to the scenes during encoding. We expected high confidence ratings and high accuracy for all scenes during this phase as a sign that participants had indeed followed instructions on all trials.

Procedure

Participants received a written invitation, which briefly explained the experiment in general terms (e.g. “you will be seeing pictures with emotional or neutral content”), while not giving away the rationale or any suggestions that it was a memory experiment. The invitation contained a warning that the pictures could be considered “frightening, unpleasant, or offensive”. Participants were also not told to which condition they had been assigned, or indeed that there were more than one condition. In the involuntary condition, they were given deliberate misinformation that the primary focus of the task was attention rather than memory. Finally, participants were not told about the recognition task before they actually had to do it.

Upon arriving at the laboratory, participants were again informed that they would see pictures that with strong negative content, that participation was completely voluntary, and that they could withdraw consent to participate at any time. They then gave their written consent and were seated in front of a computer with a 19” monitor set at 1280 x 1024 resolution. The computer also had headphones, a keyboard, and a response box attached.

They then completed the three phases of the task in the fixed order described above. Likert-style responses were given using the response box with the numbers 1 to 5 clearly labelled on the buttons. When participants recorded a retrieved scene, they would use the computer's keyboard to write the keywords. After completing the recognition phase, participants watched 10 positively valenced pictures from the IAPS in order to induce a positive mood, were debriefed, and given two cinema tickets for their assistance.

Data preparation

Errors in the retrieval phase were defined in the following way: First, if participants in the involuntary condition recorded a memory of a scene in response to a filler sound - or if participants in either condition described a scene that could not be readily identified - this was recorded as an error. Second, if participants in either condition retrieved a scene in response to a sound different from the one it had been paired with during encoding, this was recorded as an indiscriminately cued memory. According to our predictions, emotional scenes should be favoured at retrieval independent of the distinctiveness of the associated cue. This could entail that the association between sound cue and scene is less important for emotional scenes compared to neutral scenes, which could increase the rate of indiscriminately cued memories. Tentatively, this would be consistent with the object-based binding framework proposed by Mather (2007) that emotional arousal reduce the encoding and binding of unrelated features, in this case the arbitrarily related sound cue.

The coding of memories was performed by the first author, and an independent judge with no knowledge of the hypotheses coded 20 % of the memories to ensure the reliability of the original coding. Interrater agreement for Experiment 1 was 100 %. We calculated the proportion of error types to the total number of recorded scenes in each condition. For errors this was: 19.94 % (involuntary condition) and 0.00 % (voluntary condition). For indiscriminately cued memories the numbers were: 9.35 % (involuntary condition) and 13.03 % (voluntary condition). Errors were omitted from all analyses, while indiscriminately cued memories were included as valid responses.

The reasons for not excluding indiscriminately cued memories were the following: 1) indiscriminately cued memories represented memories of encoded scenes – in this sense they were valid memories, albeit retrieved in response to a different sound than the one with which they had been presented during encoding. 2) In the involuntary condition, participants were not asked to monitor for accuracy of the sound-scene match – which would be counter to the concept of involuntary retrieval – but rather to indicate if *any* scene came to mind. 3) Naturally, it did not make sense to distinguish between subjectively highly similar – although objectively different – repeated cues in our coding scheme. As a result of this, the two conditions with repeated cues showed a much lower rate of indiscriminately cued memories compared with conditions with unique cues, also speaking against excluding such cases. 4) It was theoretically relevant to examine whether the emotional scenes would result in more indiscriminately cued memories than the neutral scenes, given that the distinctiveness of the sound cue in relation to the scene would be less decisive for their activation, according to our predictions.

Results

Demographic Variables

Two independent *t*-tests confirmed that participants in the two conditions did not differ significantly on age (involuntary: *M* = 22.43, *SD* = 1.79; voluntary: *M* = 22.19, *SD* =

4.62, *t*(30) = 0.20, *p* = .841), or gender (involuntary: 25 % male, voluntary: 19 % male, *t*(30) = 0.42, *p* = .681).

Ratings Obtained at Encoding

We analysed ratings of intensity and relatedness in the encoding phase by means of 2 (Sound: unique *vs.* repeated sound) x 2 (Emotion: emotional *vs.* neutral scene) repeated measures ANOVAs. As expected, emotional scenes were rated as more emotionally intense than neutral pictures (*M* = 4.53, *SD* = 0.42 vs. *M* = 1.56, *SD* = 0.29; main effect of Emotion: $F(1,31) = 1665.65$, $p < .0001$, $\eta_p^2 = .98$). The sound cues did not show any effect or interaction with emotion in relation to ratings of intensity (*p*s > .105). Emotional scenes were rated as less related to their paired sounds compared with neutral scenes (main effect of Emotion: $F(1,31) = 19.56$, $p < .0001$, $\eta_p^2 = .39$), with no other effects being statistically significant (*p*s > .186).

Involuntary and Voluntary Memories across the Four Conditions

In order to analyse the proportion of retrieved memories in relation to the total number of cues in each trial category (e.g., UniCue_EmoScene), we conducted a 2 (Cueing: repeated *vs.* unique sounds) x 2 (Emotion: emotional *vs.* neutral scenes) x 2 (Retrieval: involuntary *vs.* voluntary) repeated measures ANOVA, with involuntary versus voluntary retrieval condition as a between-groups variable and Cueing and Emotion as repeated measures factors. The findings are illustrated in Figure 1. As expected, participants in the voluntary condition recalled proportionally more scenes compared with participants in the involuntary condition. Also, participants overall retrieved more scenes in response to unique relative to repeated cues (main effect of Cueing: $F(1,30) = 36.43$, $p < .001$, $\eta_p^2 = .55$; main effect of Retrieval: $F(1,30) = 37.36$, $p < .0001$, $\eta_p^2 = .56$; see Figure 1). The main effect of Emotion was not significant ($F(1,30) = 0.48$, $p = .490$, $\eta_p^2 = .02$), and Cueing and Emotion did not interact ($F(1,30) = 1.15$, $p = .292$, $\eta_p^2 = .04$). No other effects were significant (*p*s > .20).

Next, we turned to the characteristics recorded for each retrieved scene. We conducted a series of repeated measures ANOVAs for each individual characteristic (intensity, bodily reactions, specificity, and reaction time). For each analysis, retrieval condition was a between-groups variable, while valence of the scenes (emotional *vs.* neutral) was a repeated measure. For this analysis we did not include scenes originally paired with a repeated sound (i.e., RepCue_EmoScene and RepCue_NeuScene) due to the fact that few participants recorded memories in response to the repeated cues, consistent with our predictions. As a consequence, there was a high number of missing values for these types of trials, which prevented us from including them in the analyses. Analysing the memories recorded in response to unique cues, responses from 12 participants in the involuntary condition and 16 participants in the voluntary condition were available. The means, standard deviations, and *F*-statistics can be seen in Table 1. We found main effects of Emotion for measures of intensity, bodily reactions, and specificity. Memories of emotional scenes were rated as more intense and with stronger bodily reactions compared with neutral scenes. Memories of neutral scenes, however, were more specific than memories of emotional scenes. Finally – as predicted - we found that participants in the voluntary condition were slower to retrieve scenes regardless of their valence.

To examine whether participants had indeed paid attention to - and where capable of remembering - both the emotional and neutral scenes, they completed a recognition task following the retrieval task. As anticipated, the results of this task showed a high mean accuracy (99 - 100 %) and confidence ratings (4.81 – 4.94) for the different types of scenes, indicating that the participants had followed instructions during encoding and were reporting memories of the previously presented scenes during the retrieval phase.

Indiscriminately cued memories

The percentage of indiscriminately cued memories in relation to the number of memories retrieved in response to unique cues was analysed with a 2 (Emotion: emotional *vs.* neutral scene) x 2 (Retrieval: voluntary *vs.* involuntary) repeated measures ANOVA. We did not include repeatedly cued memories in the analysis for reasons given earlier. Participants had more indiscriminately cued memories for emotional scenes compared with neutral scenes (22.3 % *vs.* 14.5 %; main effect of Emotion: $F(1,30) = 5.32$, $p = .028$, $\eta_p^2 = .15$; see Table 2). No other effects were significant (*p*s > .40).

Discussion

We found partial support for our initial hypotheses. Following the prediction from cue-item discriminability, the frequency of memories was greatest in response to unique cues in both retrieval conditions. However, inconsistent with our predictions, emotion did not appear to influence the accessibility of the scenes.

As expected, memories of emotional scenes were rated as more intense and accompanied by stronger bodily reactions compared with memories of neutral scenes, while the latter were rated as more specific. Emotional scenes also led to more indiscriminately cued memories. Finally, as hypothesized, voluntary retrieval was slower than involuntary retrieval, but it also generated a greater proportion of memories in response to cues.

What are the possible explanations for the lack of an effect of emotion on retrieval frequency? One possibility might be that effects of emotion will be seen only with greater retention intervals, which would allow more time for consolidation and rehearsal processes to influence accessibility. Before testing the prediction that increasing the retention interval will selectively increase the accessibility of emotional scenes, a few issues with the present design should be considered. First, it could be argued that the emotional scenes were not sufficiently intense to allow the expected effects to be observed. However, we find this explanation unlikely. Participants rated the intensity of the emotional scenes well above 4 on a 5-point scale, several of them spontaneously remarked how unpleasant the scenes were, and a few participants were mildly distressed by the procedure. Second, the perceived strength of the relatedness between the scenes and the sounds during encoding could confound the results. We found that ratings of relatedness were significantly lower for emotional scenes compared with neutral scenes. If participants perceived the associative strength of the emotional pairs as poorer than that of the neutral pairs, this might explain why they reported equal numbers of emotional and neutral memories. In addition, it could explain why emotional scenes led to more indiscriminately cued memories. In order to rule out this potential confounder, we ran a replication of Experiment 1 in which we controlled the relatedness of the stimulus-pairs instead of using a randomized procedure.

Experiment 2

Method

Thirty-two participants (28 female, mean age 22.72, range 19-32) were recruited from Aarhus University's student population and randomly assigned to either voluntary or involuntary retrieval. The design, material, method, and procedure were identical to

Experiment 1, except for one important modification: Based on the ratings of relatedness between individual cues and scenes in Experiment 1, we constructed a set of cue-scene pairs that had equal ratings of relatedness for emotional and neutral scenes. This meant that the stimuli were no longer paired randomly. We did so in order to avoid any biases due to variations in perceived relatedness across the four conditions. Errors and indiscriminately cued memories were coded as in Experiment 1. Again an independent judge coded ~20% of the memories in order to ensure reliability. Interrater agreement was 90 %. In cases of disagreement, we compared the two codings and decided on the most appropriate one. This generally meant that the original coding was retained. The error rates were 26.59 % (involuntary condition) and 1.49 % (voluntary condition). For indiscriminately cued memories the numbers were 9.38 % (involuntary condition) and 13.28 % (voluntary condition). Again, errors were omitted from the analyses while indiscriminately cued memories were included for the reasons provided in Experiment 1.

Results

Demographics

Independent *t*-tests confirmed that there were no statistically significant differences between participants in the voluntary and involuntary condition in relation to age (involuntary: M = 22.00, SD = 1.63; voluntary: M = 23.44, SD = 3.44, $t(30)$ = 1.51, p = .142), and gender (involuntary: 6 % male, voluntary: 19 % male, $t(25.05)$ = 1.05, p = .302).

Ratings Obtained at Encoding

As in Experiment 1, we analysed ratings of intensity and relatedness in the encoding phase by means of 2 (unique *vs.* repeated sound) x 2 (emotional *vs.* neutral scene) repeated measures ANOVAs. We again found higher ratings of intensity for emotional compared with neutral scenes (main effect of Emotion: $F(1,31) = 838.32$, $p < .0001$, $\eta_p^2 = .96$). In contrast to Experiment 1, there was no effect of emotion on relatedness ($F(1,31) = 0.75$, $p = .393$, $\eta_p^2 = .02$), however, participants had lower ratings of relatedness for scenes paired with unique sounds compared with scenes paired with repeated sounds (main effect of Sound: $F(1,31) = 13.13$, $p = .001$, $\eta_p^2 = .30$). All other effects in these analyses were statistically non-significant (*p*s > .28).

Involuntary and Voluntary Memories across the Four Conditions

As in Experiment 1, we conducted a 2 (Cueing: repeated *vs.* unique sounds) x 2 (Emotion: emotional *vs.* neutral scenes) x 2 (Retrieval: involuntary *vs.* voluntary) repeated measures ANOVA, with involuntary versus voluntary retrieval condition as a between-groups variable and Cueing and Emotion as repeated measures. The findings are illustrated in Figure 2. As in Experiment 1, participants in the voluntary condition recalled proportionally more scenes compared with participants in the involuntary condition, and all participants retrieved more scenes in response to unique relative to repeated cues (main effect of Cueing: $F(1,30) = 37.64$, $p < .0001$, $\eta_p^2 = .56$; main effect of Retrieval: $F(1,30) = 31.50$, $p < .0001$, $\eta_p^2 = .51$; see Figure 2). As in Experiment 1, the main effect of Emotion was not significant ($F(1,30) = 1.87$, $p = .182$, $\eta_p^2 = .06$). We found an interaction between Emotion and Cueing ($F(1,30) = 11.43$, $p = .002$, $\eta_p^2 = .28$). Surprisingly, the interaction was due to a higher frequency of *neutral* scenes in response to unique cues ($t(31) = 3.91$, $p < .001$), counter to the prediction of an advantage of emotional scenes and an additive effect of emotion and cueing. Repeated cues did not show this advantage of neutral scenes ($t(31) = 0.78$, $p = .442$). Finally, we found a greater advantage of voluntary retrieval when memories were cued by unique sounds compared with memories that were cued by repeated sounds (interaction between Cueing and Retrieval: $F(1,30) = 5.65$, $p = .024$, $\eta_p^2 = .16$; see Figure 2). There was no significant interaction between Emotion and Retrieval ($F(1,30) = 0.01$, $p = .934$, $\eta_p^2 = .00$).

For the analysis of the characteristics of retrieved scenes, we again excluded memories to repeated cues (i.e., RepCue_EmoScene and RepCue_NeuScene) due to the number of missing values for these types of trials. Data from 15 participants in the involuntary condition and 16 participants in the voluntary condition were available for the analyses. The means, standard deviations, and *F*-statistics can be seen in Table 1. We replicated the findings from Experiment 1 that memories of emotional scenes were rated as more intense and with stronger bodily reactions than memories of neutral scenes, while neutral memories had higher ratings of specificity. In addition, participants in the voluntary condition were slower to retrieve scenes compared with participants in the involuntary condition.

Again, the recognition task showed a high accuracy (98 – 100 %) and high confidence ratings (4.78 – 4.96) for the different types of scenes, indicating that they were successfully encoded, and that the participants were reporting actual memories during the retrieval phase.

Indiscriminately cued memories

The proportion of indiscriminately cued memories was again analysed with a 2 (emotional *vs.* neutral scenes) x 2 (involuntary *vs.* voluntary retrieval) repeated measures ANOVA. No effects were significant (*ps* > .403; see Table 2).

Discussion

We replicated the finding from Experiment 1 that scenes associated with unique cues led to more memories in both retrieval conditions. We also replicated the finding that emotion did not seem to influence the frequency with which the scenes were remembered, since overall memory rates of emotional and neutral scenes were the same. In addition, emotional scenes cued by unique sounds were actually retrieved with reduced frequency compared with neutral scenes cued by unique sounds. This finding differed from Experiment 1, where uniquely cued emotional and neutral scenes were retrieved equally often. Since the stimuli were identical in the two experiments, it is possible that differences in the two samples of participants were responsible for this discrepancy. While the age and gender distribution was equal in the two samples (age: $t(62) = 0.99$, $p = .328$; gender: $t(62) = 0.52$, $p = .604$), it is possible that other factors, such as state anxiety or other measures of negative affect, differed between the two samples, which in turn may have influenced the results. Irrespective of these possibilities, Experiment 2 replicated the key finding that memories of emotional scenes were not recorded more frequently than neutral scenes.

Since there was no difference in ratings of relatedness between emotional and neutral scenes in the encoding phase of Experiment 2, it is unlikely that the perceived strength of the associations was responsible for the lack of an effect of emotion. We also replicated the findings from Experiment 1 that memories of emotional scenes were rated as more intense and lead to stronger bodily reactions than memories of neutral scenes, but the latter were rated as more specific. The rates of indiscriminately cued memories did not differ between neutral and emotional scenes in contrast to findings from Experiment 1. Finally, voluntary retrieval was slower than involuntary retrieval, but generated more memories, similar to findings in Experiment 1.

In summary, we successfully replicated all the main findings of Experiment 1 when controlling for relatedness between emotional and neutral pairs. We therefore feel confident in considering theoretical explanations for the lack of an effect of emotion on retrieval frequency. It could be hypothesized that it is not the memories of the emotional scenes that show a reduced accessibility, but rather the memories of the neutral scenes that show an inflated accessibility relative to what might be found with a longer retention interval. A number of studies examined the interaction between emotional arousal during encoding of paired associates and retention interval as part of an effort to experimentally investigate the theory of Freudian repression (Bradley & Baddeley, 1990; Kleinsmith & Kaplan, 1963, 1964; Kleinsmith, Kaplan, & Tarte, 1963; Levinger & Clark, 1961; Levonian, 1967; Parkin, Lewinson, & Folkard, 1982). These early studies found that immediate testing led to fewer forgotten associations when arousal was low compared to when arousal was high, whereas longer retention intervals reversed this effect. However, later studies did not find any effects of emotion on immediate memory tests (Kebeck & Lohaus, 1986; Sharot & Yonelinas, 2008; Talmi & McGarry, 2012) or found superior memory for emotional stimuli (see Ferré, 2002, for a review). This pattern of findings suggests that while emotional stimuli are better remembered over longer retention intervals, their effect on immediate memory is much less consistent. One possible explanation might be that the effects of neurophysiological consolidation, rehearsal, and interference require longer retention intervals to exert an effect. While estimates vary depending on the nature of the stimulus, in most cases consolidation is not completed until several hours following the experience (Dudai, 2004). Some research indicates that sleep – particularly dream sleep – reinforces consolidation of emotional memories (Crick & Mitchison, 1983; LaBar & Cabeza, 2006; Stickgold, Hobson, Fosse, & Fosse, 2001; Wagner, Gais, & Born, 2001). At another level of analysis, rehearsal and interference processes would obviously need time to assert their influence on memory (e.g., Bjork, 2003; Roediger, Weinstein, & Agarwal, 2010).

In order to test the hypothesis that the apparent failure of emotion to enhance memory accessibility was due to the very short retention interval, we conducted Experiment 3, where the encoding and retrieval phases were separated by a 24-hour interval.

Experiment 3

Method

Thirty-two participants (23 female, mean age 23.16, range 20-35) were recruited from Aarhus University's student population and randomized to either voluntary or involuntary retrieval. The design, material, method, and procedure were identical to Experiment 2, except for one important modification: After participants had completed the encoding task, they left the lab and came back 24 hours later for the retrieval task, recognition task, and debriefing (given the instructions, they were not told that they would be retrieving the scenes during this second visit).

Errors and indiscriminately cued memories were coded as in Experiment 1. Again, an independent rater coded ~20% of the memories. Interrater agreement was 95 %. As in Experiment 2, we re-evaluated disputed memories and changed the original coding if appropriate. The error rates were 23.20 % (involuntary condition) and 2.08 % (voluntary condition). For indiscriminately cued memories the numbers were 29.73 % (involuntary condition) and 25.53 % (voluntary condition). Errors were omitted from the analyses while indiscriminately cued memories were included.

Results

Demographics

Participants in the voluntary and involuntary condition were comparable in relation to age (involuntary: M = 24.06, SD = 4.61; voluntary: M = 22.25, SD = 1.44, $t(16.30) = 1.44$, $p = .168$) and gender (involuntary: 37 % male, voluntary: 19 % male, $t(28.71) = 1.17$, $p = .253$).

Ratings Obtained at Encoding

We again analysed ratings of intensity and relatedness in the encoding phase by means of 2 (Sound: unique *vs.* repeated sound) x 2 (Emotion: emotional *vs.* neutral scene) repeated measures ANOVAs. As in Experiments 1 and 2, participants rated the emotional scenes as more intense than the neutral scenes (main effect of Emotion: $F(1,31) = 623.37$, $p < .0001$, $\eta_p^2 = .95$). In addition, we also found an interaction between Emotion and Sound ($F(1,31) = 8.14$, $p = .008$, $\eta_p^2 = .21$) in that emotional scenes paired with unique sounds were rated as more intense than emotional scenes paired with repeated sounds ($t(31) = 2.42$, $p = .021$), with no statistically significant difference between neutral scenes ($t(31) = -0.75$, $p = .460$). The main effect of Sound was not significant ($F(1,31) = 1.91$, $p = .177$, $\eta_p^2 = .06$). Concerning relatedness, we found lower ratings for scenes paired with unique sounds compared with scenes paired with repeated sounds, and emotional scenes compared with neutral scenes (main effect of Emotion: $F(1,31) = 11.09$, $p = .002$, $\eta_p^2 = .26$; main effect of Sound: $F(1,31) = 6.22$, $p = .018$, $\eta_p^2 = .17$). The interaction between Emotion and Sound was not significant ($F(1,31) = 0.22$, $p = .643$, $\eta_p^2 = .01$).

Involuntary and Voluntary Memories across the Four Conditions

As in the previous experiments, we conducted a 2 (Cueing: repeated *vs.* unique sounds) x 2 (Emotion: emotional *vs.* neutral scenes) x 2 (Retrieval: involuntary *vs.* voluntary) repeated measures ANOVA, with involuntary versus voluntary retrieval condition as a between-groups variable and Cueing and Emotion as repeated measures. The findings are presented in Figure 3. Participants in the voluntary condition recalled proportionally more scenes compared with participants in the involuntary condition, and participants in both conditions retrieved more memories of emotional than of neutral scenes (main effect of Emotion: $F(1,30) = 10.18$, $p = .003$, $\eta_p^2 = .25$; main effect of Retrieval: $F(1,30) = 85.46$, $p < .0001$, $\eta_p^2 = .74$; see Figure 3). The main effect of Cueing was not significant ($F(1,30) = 0.70$, $p = .411$, $\eta_p^2 = .02$), and Cueing and Emotion did not interact ($F(1,30) = 2.80$, $p = .104$, $\eta_p^2 = .09$).No other effects were significant (all *ps* > .154).

For the analysis of the characteristics of retrieved scenes, we again excluded memories in response to repeated cues (i.e., RepCue_EmoScene and RepCue_NeuScene) due to the number of missing values for these types of trials. Data from 15 participants in the involuntary condition and 14 participants in the voluntary condition were available for the analyses. The means, standard deviations, and *F*-statistics can be seen in Table 3. As in the previous experiments, memories of emotional scenes were rated as more intense and with stronger bodily reactions than memories of neutral scenes, while memories of neutral scenes had higher ratings of specificity. Participants in the voluntary condition had slightly higher ratings of intensity for the memories of emotional scenes as compared to participants in the involuntary condition, but slightly lower ratings for the memories of neutral scenes, which created a significant interaction. Finally, participants in the voluntary condition again had longer retrieval times compared with participants in the involuntary condition.

As in previous experiments, post-retrieval recognition showed a high accuracy (98 – 100 %) and confidence (4.47 – 4.85) across emotional and neutral scenes.

Indiscriminately cued memories

Proportions of indiscriminately cued memories in response to unique cues were analysed with a 2 (Emotion: emotional *vs.* neutral scenes) x 2 (Retrieval: involuntary *vs.* voluntary retrieval) repeated measures ANOVA. Participants generated more indiscriminately cued memories for emotional than neutral scenes (43.7 % *vs.* 29.6 %; main effect of Emotion: $F(1,30) = 5.55$, $p = .025$, $\eta_p^2 = .16$; see Table 2).

Discussion

Compared with the findings in Experiments 1 and 2, Experiment 3 reversed the effects of cueing and emotion. In accordance with theories of consolidation and interference, we found that participants retrieved higher frequencies of emotional memories. In contrast to Experiments 1 and 2, memory frequency was unrelated to the uniqueness of the associated cues. This finding leaves us with a dissociation between the effect of cueing and the effect of emotion that appears to be time-dependent, given the fact that the effect of emotion was absent in Experiments 1 and 2 with very short intervals separating encoding and retrieval. Intriguingly, this may suggest a mechanism capable of explaining intrusive memories of traumatic events in response to highly frequent cues as described in clinical anecdotes: The accessibility of highly emotional events is enhanced and becomes less cue-dependent with the passage of time. We will return to this possibility in the General Discussion.

We again found that memories of emotional scenes were rated as more intense and with stronger bodily reactions, while memories of neutral scenes were more specific.

Voluntary retrieval was slower but generated more memories than did involuntary retrieval. Finally, participants reported a greater number of indiscriminately cued memories for emotional scenes than for neutral scenes.

Although we used the exact same pairing of scenes and sounds as in Experiment 2, this time ratings of relatedness showed effects for both cueing (as in Experiment 2) and emotion (as in Experiment 1). This suggests that perceived relatedness reflects individual differences between participants that may or may not be related to retrieval success. In order to explore this idea further, we calculated each participant's relatedness score for each individual trial as well as the retrieval frequency for the cue belonging to that trial. We then paired up relatedness scores with retrieval frequencies for each individual trial and correlated those data for each participant. Out of the 32 resulting correlations, only three were significant at the $p < .05$ level and 14 were in fact negative. We transformed each correlation into a z-score in order to calculate an average correlation. This was $r = .059$. We therefore conclude that ratings of relatedness during encoding had no systematic influence on the retrieval of memories in Experiment 3, consistent with the fact that the results were in the opposite direction of what the scores for relatedness would predict.

When considering our main finding that emotion – and not cueing – determined retrieval frequency after a retention interval of 24 hours, an important question arises: Why does it take time for the emotional memories to gain their expected advantage over the neutral memories? In earlier research investigating the effect of retention interval on memory, the superiority of emotional stimuli following a delay has been interpreted in terms of a differing pattern of optimal arousal (e.g. Kleinsmith & Kaplan, 1963; Parkin et al., 1982). At immediate testing, the high arousal is detrimental to memory, but over a longer delay memory is enhanced by high arousal during encoding. The theory of consolidation reviewed in the introduction offers a possible neurophysiological account of this mechanism.

Based on the theory of consolidation as well as behavioural research on mechanisms of forgetting (e.g., Roediger et al., 2010), we would expect interference to be more pronounced for neutral memories, making them more prone to forgetting than emotional memories over time. This could explain why their retrieval frequency appears to drop more sharply compared with that of emotional memories – their accessibility is reduced more quickly, since they are less efficiently consolidated. While we cannot make any firm conclusion regarding the detailed timeline of when this consolidation is completed, our results are consistent with the animal literature showing that the first stage of consolidation takes hours at least (Dudai, 2004).

Before discussing the effect of time on involuntary memories of emotional and neutral scenes further, we wanted to replicate the findings from Experiment 3, but this time extending the retention interval to a week. One important reason behind this choice was the distinction between short-term and long-term consolidation mentioned earlier. While short-term consolidation takes up to a day to complete, long-term consolidation lasts days or even months (Dudai, 2004). This suggest that the emotional scenes should be accessible to voluntary and involuntary retrieval even after a week, in spite of the fact that encoding was incidental rather than goal-driven, that the scenes were unlikely to have strong personal relevance, and that each scene was presented only briefly.

Experiment 4

Method

Thirty-two participants (26 female, mean age 24.62, range 20-63) were recruited from Aarhus University's student population and randomly assigned to either the voluntary or involuntary condition. The design, materials, methods, and procedure in Experiment 4 were identical to Experiment 3, except for the following changes: 1) The retention interval was extended to seven days instead of one day. 2) By now we felt confident that relatedness did not influence the retrieval of memories and since the fixed pairings of Experiments 2 and 3 did not give consistent results, we returned to the random pairings from Experiment 1.

Errors in the retrieval phase were coded in the same way as in Experiments 1-3. Interrater agreement was 97 %. We calculated the proportion of error types to the total number of retrieved scenes. For errors this was 16.94 % (involuntary condition) and 1.86 % (voluntary condition); for indiscriminately cued memories: 32.95 % (involuntary condition) and 34.10 % (voluntary condition). Errors were omitted, and indiscriminately cued memories were retained, in all analyses of memory frequency.

Results

Demographics

Independent *t*-tests confirmed that there were no statistically significant differences between participants in the voluntary and involuntary condition in relation to age (involuntary: M = 25.86, SD = 11.19; voluntary: M = 23.47, SD = 2.95, $t(30) = 0.80$, p = .431) and gender (involuntary: 19 % male, voluntary: 19 % male).

Ratings Obtained at Encoding

We again analysed ratings of intensity and relatedness in the encoding phase by means of 2 (Sound: unique *vs.* repeated sound) x 2 (Emotion: emotional *vs.* neutral scene) repeated measures ANOVAs. For ratings of intensity, emotional scenes received higher ratings than neutral scenes (M = 4.53, SD = 0.45 *vs.* M = 1.69, SD = 0.34; main effect of Emotion: $F(1,31) = 1022.76$, $p < .0001$, $\eta_p^2 = .97$). Scenes paired with unique sounds received a slightly higher rating compared with scenes paired with sounds that were repeated during encoding (M = 3.14, SD = 0.32 *vs.* M = 3.07, SD = 0.33; main effect of Sound: $F(1,31) = 4.36$, $p = .045$, $\eta_p^2 = .12$). The interaction between Emotion and Sound was not significant ($F(1,31) = 2.04$, $p = .164$, $\eta_p^2 = .06$). Ratings of relatedness showed no significant effects or interactions (all *p*s > .168).

Involuntary and Voluntary Memories across the Four Conditions

As in the previous studies, we conducted a 2 (Cueing: repeated *vs.* unique sounds) x 2 (Emotion: emotional *vs.* neutral scenes) x 2 (Retrieval: involuntary *vs.* voluntary) repeated measures ANOVA, with involuntary versus voluntary retrieval condition as a between-groups variable and Cueing and Emotion as repeated measures. The findings are shown in Figure 4. Participants retrieved more memories for emotional than for neutral scenes. Participants in the voluntary condition retrieved proportionally more memories overall compared with participants in the involuntary condition, and this advantage was greatest for the emotional scenes (main effect of Emotion: $F(1,30) = 18.44$, $p < .001$, $\eta_p^2 = .38$; main effect of Retrieval: $F(1,30) = 38.78$, $p < .0001$, $\eta_p^2 = .56$; interaction between Emotion and Retrieval: $F(1,30) = 7.64$, $p = .010$, $\eta_p^2 = .20$; see Figure 4). However, the difference between emotional and neutral scenes was still significant during involuntary retrieval (emotional: M

= 18.95, *SD* = 12.98; neutral: *M* = 11.33, *SD* = 12.34; main effect of Emotion: $F(1,15) = 8.92$, $p = .009$, $\eta_p^2 = .37$; main effect of Cueing: $F(1,15) = 5.35$, $p = .035$, $\eta_p^2 = .26$). As in Experiment 3, the main effect of Cueing was not significant ($F(1,30) = 1.96$, $p = .172$, $\eta_p^2 = .06$), and Cueing and Emotion did not interact ($F(1,30) = 0.55$, $p = .465$, $\eta_p^2 = .02$). No other effects were significant (all *p*s > .395).

Contrary to Experiments 1, 2, and 3, there were fewer missing values in the repeated cue conditions compared with the unique cue conditions. As a result, we included only scenes originally paired with a repeated sound (i.e., RepCue_EmoItem and RepCue_NeuItem) in the analyses on subjectively rated memory qualities. This left us with 11 participants in the involuntary condition and 13 participants in the voluntary condition. The means, standard deviations, and *F*-statistics can be seen in Table 3. We found main effects of Emotion for ratings on intensity and bodily reactions. Memories for emotional scenes were rated as more intense and with stronger bodily reactions, compared with memories for neutral scenes. We also found that participants in the voluntary condition were slower to retrieve memories.

Recognition showed high accuracy (96 – 99 %) and confidence (4.10 – 4.67) for both emotional and neutral scenes after a week, again indicating that the participants were reporting actual memories during the retrieval phase.

Indiscriminately cued memories

Proportions of indiscriminately cued memories were analysed with a 2 (Emotion: emotional *vs.* neutral scenes) x 2 (Retrieval: involuntary *vs.* voluntary retrieval) repeated measures ANOVA. As in Experiments 1 and 3, participants generated more indiscriminately cued memories for emotional scenes compared with neutral scenes, however, this was only a near-significant trend in the analysis (52.1 % *vs.* 35.9 %; main effect of Emotion: $F(1,30) = 4.06$, $p = .053$, $\eta_p^2 = .12$; see Table 2).

Discussion

We replicated the findings from Experiment 3 after a week-long retention interval. Emotional scenes were more frequently remembered than neutral scenes, and uniqueness of the cues did not show an effect. This time, the randomization of sound cues and scenes worked in the sense that ratings of relatedness showed only statistically non-significant effects. This further strengthens the observations in Experiments 1-3 that the perceived relatedness did not impact retrieval frequency. Memories of emotional scenes were rated as more intense and involved stronger bodily reactions at the time of recall than memories of neutral scenes. As in all previous experiments, voluntary retrieval was slower and led to more memories than involuntary retrieval. Finally, indiscriminately cued memories were higher for memories of emotional scenes compared with memories of neutral scenes, consistent with findings from Experiments 1 and 3.

The results of the four experiments suggest that there is an interaction between emotion, cueing and the passage of time. Since all four experiments were conducted in the same laboratory with the same materials and equipment, and since all subjects were recruited from the same population of students and none of them participated in more than one of the four studies, it was possible to reanalyse the data with the specific aim of testing the interaction between emotion, cueing and time.

Meta-Analysis of Experiments 1 to 4

Method

We included data from all 128 participants from the previous four experiments in one database. Participant group membership in relation to time interval was defined by the retention interval between encoding and retrieval. This entailed that participants from Experiments 1 and 2 were assigned to one group with a short retention interval of a few minutes, while participants from Experiment 3 and 4 were assigned to individual groups with intervals of one day and one week, respectively.

Results

Involuntary and Voluntary Memories across the Four Conditions

In the meta-analysis of the previous results, we conducted a 2 (Cueing: repeated *vs.* unique sounds) x 2 (Emotion: emotional *vs.* neutral scenes) x 2 (Retrieval: involuntary *vs.* voluntary) x 3 (Interval: short *vs.* one day *vs.* one week) repeated measures ANOVA, with Retrieval and Interval as between-groups variables, and Cueing and Emotion as repeated measures. Since we were running an additional analysis on the same dataset, we adjusted the alpha-level to $\alpha = .01$ in order to reduce the risk of type-1 error.

The findings from the meta-analysis are represented in Figures 5 and 6. We chose to represent the effects of Emotion x Time x Retrieval and Cueing x Time x Retrieval separately since there was no significant 4-way interaction involving all factors in the model ($F(2,122) = 0.81$, $p = .923$, $\eta_p^2 < .01$). Before examining these 3 way interaction effects in greater detail, two main effects of the analysis were important: Memories showed forgetting over time and voluntary retrieval generated a greater proportion of memories compared with involuntary retrieval (main effect of Interval: $F(1,122) = 5.96$, $p = .003$, $\eta_p^2 = .09$; main effect of

Retrieval: $F(1,122) = 145.07$, $p < .001$, $\eta_p^2 = .54$). We also found that – overall – repeatedly cued memories of neutral scenes were less frequent than the other three types of memories (interaction between Emotion and Cueing ($F(1,122) = 8.11$, $p = .005$, $\eta_p^2 = .06$).

As can be seen in Figure 5, memories of emotional scenes were more frequent in both retrieval conditions, which was shown by a main effect of Emotion ($F(1,122) = 25.48$, $p < .0001$, $\eta_p^2 = .17$). Most importantly, frequency of memories for emotional scenes did not decrease over the three time intervals (short: $M = 40.08$, $SD = 27.48$; one day: $M = 47.46$, $SD = 26.03$; one week: $M = 41.11$, $SD = 33.17$), whereas memories for neutral scenes did decrease in frequency over time (short: $M = 41.65$, $SD = 24.53$; one day: $M = 32.42$, $SD = 20.56$; one week: $M = 19.73$, $SD = 15.40$) creating an interaction between Emotion and Interval ($F(1,122) = 10.78$, $p < .001$, $\eta_p^2 = .14$). Finally, there was a 3-way interaction between Emotion, Interval, and Retrieval ($F(1,122) = 5.01$, $p = .008$, $\eta_p^2 = .08$). This effect reflected the finding from Experiment 4 that the difference between emotional memories and neutral memories was greater during voluntary retrieval compared with involuntary retrieval after a 1-week interval, while the interactions between Emotion and Retrieval were not significant after the short and 1-day intervals (short: $F(1,62) = 1.16$, $p = .286$, $\eta_p^2 = .02$; 1-day: $F(1,30) = 1.25$, $p = .272$, $\eta_p^2 = .04$; see Figure 5).

Finally, Emotion did not significantly interact with Retrieval given our adjusted alpha level ($F(1,122) = 5.44$, $p = .021$, $\eta_p^2 = .04$). However, since our prediction was that the two retrieval conditions should not differ with regard to emotional memories, we conducted follow-up analyses to tease apart this statistical trend. It appears that it was caused by a greater difference between the frequency of emotional and neutral memories during voluntary retrieval (mean difference = 11.62, $SD = 35.30$) compared with involuntary retrieval (mean difference = 5.03, *SD* = 15.57), however the overall pattern was identical in the two retrieval conditions (see Figure 5).

Turning now to Figure 6, we found a greater frequency of memories in response to unique cues overall and greater forgetting of uniquely cued scenes over time (main effect of Cueing: $F(1,122) = 11.10$, $p = .001$, $\eta_p^2 = .08$; interaction between Cueing and Interval: $F(1,122) = 19.30$, $p < .001$, $\eta_p^2 = .24$). Finally, given our adjusted alpha level, there was a non-significant trend for a 3-way interaction between Cuing, Interval, and Retrieval ($F(2,122) = 3.91$, $p = .023$, $\eta_p^2 = .06$). This effect was due to a greater drop for uniquely cued memories in the involuntary condition after 1 week (see Figure 6). No other effects were significant (*p*s > .15). The raw frequencies for retrieval of memories across all four experiments can be seen in the Appendix.

Indiscriminately cued memories

Proportions of indiscriminately cued memories were analysed with a 2 (Emotion: emotional *vs.* neutral scenes) x 2 (Retrieval: involuntary *vs.* voluntary retrieval) x 3 (Interval: short *vs.* one day *vs.* one week) repeated measures ANOVA. The analysis showed a higher proportion of indiscriminately cued memories for emotional scenes compared with neutral scenes (emotional: *M* = 33.7 %, *SD* = 31.7 %; neutral: *M* = 24.2 %, *SD* = 29.7 %; main effect of Emotion: $F(1,122) = 5.98$, $p = .016$, $\eta_p^2 = .05$). We also found a general increase in indiscriminately cued memories over time (main effect of Interval: $F(1,122) = 18.47$, $p < .0001$, $\eta_p^2 = .23$; see Table 2). No other effects were statistically significant (all *ps* > .322).

Discussion

In the meta-analysis of Experiments 1 to 4 we confirmed our prediction that emotional scenes interact with retention interval to increase the availability of memories compared with neutral scenes. The frequency of emotional scenes did not interact with retrieval condition, but there was a significant 3-way interaction between emotional scenes, retention interval, and retrieval condition. This effect was due to a larger drop in the frequency of memories for emotional scenes in the involuntary compared with the voluntary condition after a week. In addition, uniquely cued scenes showed greater forgetting over time, and again the pattern was largely similar in the two retrieval conditions, although the frequency of uniquely cued memories showed a faster decline over time in the involuntary condition.

Analysis of indiscriminately cued memories showed that the frequency of these memories increased as the retention interval increased, and that they were more frequent for memories of emotional scenes.

General Discussion

Emotionally intense memories are persistent over time, vividly recalled, and sometimes unwanted and interfering with normal functioning. Such intrusive memories of emotional events are experienced as repetitive (e.g., Berntsen & Rubin, 2008) and often seem to come to mind even in response to rather vague cues. Despite their importance in psychological disorders, there currently is no broadly accepted explanation for the mechanisms governing their activation. Here we developed a paradigm that allowed us to investigate the combined effects of cue discriminability and emotion on involuntary memory.

Our findings suggest that the interplay between these two factors across time may be central to the development of intrusive memories of stressful or traumatic events.

Taken together, our four experiments and the meta-analysis showed that the effect of cue discriminability impacted retrieval minutes after encoding while the effect of emotion was seen only after a considerably longer delay. When retention intervals were very brief (Experiments 1 and 2), memories of emotional and neutral scenes showed equal likelihood of involuntary retrieval, while unique cues compared with repeated cues led to a higher number of memories. When we increased the retention interval to a day (Experiment 3) and a week (Experiment 4), two effects were observed. First, memories of emotional scenes were more frequent than memories of neutral scenes. Second, unique cues no longer led to more memories than repeated cues.

The finding that emotional scenes became more accessible compared with neutral scenes when retention intervals were increased is compatible with the early research on memory for associations between cues and neutral versus emotional words (Kleinsmith & Kaplan, 1963; Levinger & Clark, 1961; Parkin et al., 1982), and with the view that the passage of time allows consolidation to differentiate between the two types of memories (Dudai, 2004; Wagner et al., 2001). It also is consistent with theories of forgetting based on interference, since we have to assume that participants were more frequently faced with neutral events and objects during the retention interval than with mutilation and other trauma. However, it is important to note that consolidation – as opposed to interference - does not presuppose any similarity or overlap between the learned events and events acquired during the retention interval (Wixted, 2004). In this view, emotional memories are simply more resistant to retroactive interference due to the increased synaptic potentiation of hippocampal neurons. Finally, the interaction between emotion and interval is also consistent with the view that increased retention intervals would allow for more rehearsal and elaboration of the emotional scenes relative to the neutral scenes and thus reduce the forgetting of the former relative to the latter (Roediger et al., 2010). This pattern was confirmed in the meta-analysis of Experiments 1 to 4.

The second finding - the effects of the unique cues diminished over time - likely reflects a general weakening of the associative strength between scenes and cues due to interference. This might impact the unique cues more than the repeated cues due to the fact that the unique cues were heard only once for four seconds during encoding, while very similar versions of the repeated cues were heard several times, which could allow the development of a schematic representation of these subjectively similar sounds. So whereas the association between a unique cue and a specific scene could still be more distinctive than the association between a repeated cue and a scene (e.g., Hunt & Worthen, 2006), chances are that the participants would forget more of the 16 dissimilar sounds compared with the two sets of eight highly similar sounds (for repeatedly cued trials, they would basically only need to remember "birdsong" and "dog barking", whereas for the uniquely cued trials they would need to remember "water dripping", "man yawning", "bell chiming", "woman sighing" to name but a few).

Taken together the findings suggest that memories of emotional events may break away from the mechanisms that govern the involuntary retrieval of more mundane events - that is, the fact that a distinct connection between a cue and an event appears to be needed for a neutral scene to come to mind spontaneously (Berntsen et al., 2013). Because of the increased accessibility of the emotional events relative to other events in memory, memories of the former may come to mind repeatedly even in response to cues that were not uniquely associated with the events during encoding. This effect is not immediately present however.

According to our findings, it takes time for the emotional material to attain this enhanced salience in memory, possibly due to consolidation and processes of interference unfolding over time.

The same mechanism can be seen as consistent with the fact that participants overall retrieved more indiscriminately cued memories of emotional scenes: Compared with neutral scenes, the emotional scenes more frequently came to mind in response to sound cues that originally had been paired with a different scene. One possible source of explanation is Mather's (2007) object-based binding framework. The central prediction from this framework is that the arousal resulting from exposure to an emotional stimulus will increase binding of features that are considered part of the same stimulus (e.g., its colour and shape) – but will decrease binding of unrelated features (e.g., an associated but semantically unrelated cue). The proposed mechanism behind this decreased binding of between-object features is the focused attention on the emotionally arousing stimulus leaving precious few resources to process accompanying features (Mather, 2007). Since the sound cues used in the present experiment were deliberately unrelated to the scenes, the predictions from object-based binding are relevant for the present data and they fit the results. The emotional arousal experienced when viewing the emotional scenes made them more available to both voluntary and involuntary retrieval, but could also have reduced the encoding of their associated sound cues, which then led to a higher number of indiscriminately cued memories for emotional scenes in both retrieval conditions.

Together, cue-item discriminability, consolidation and interference, and object-binding can explain the findings of the present experiments. These mechanisms suggest an explanatory model of intrusive, involuntary memory for emotionally intense events. In describing this model, we consider the encoding phase of the experiments to be an emotionally intense event for the participants. The individual trial pairs are discrete details of that event. Minutes after the event, memory for the details is relatively unaffected by emotional intensity and instead memory accessibility is related to the distinctiveness of the cues. However, after a longer delay, the distinctiveness of the cues becomes less important and instead the emotional intensity associated with the scenes predicts accessibility. In the present work, this effect persisted after a week. Due to the arousal experienced during encoding, the weakening of association with unrelated details could be especially pronounced for the emotional material. This decreased binding leads to higher rates of indiscriminately cued memories for the emotional scenes.

The model proposed above does not have differential predictions for voluntary and involuntary retrieval except that the two types of memory are differentiated by the effort they involve and the resulting frequency of memories. Voluntary retrieval is slower, likely reflecting a deliberate search through the possible scenes associated with a given cue. While this strategy takes time, it also leads to a higher frequency of memories without generating more errors or indiscriminately cued memories in the process, making it overall more effective. These predictions were supported here. The findings also were consistent with the idea that voluntary and involuntary retrieval are similarly affected by emotion at encoding. This is a central tenet of the framework for autobiographical memory in relation to PTSD symptoms, proposed by Berntsen and Rubin (2008; see also Berntsen, 2009). In a recent functional magnetic resonance imaging (fMRI) study by Bourne, Mackay, and Holmes (2013), healthy adults were shown traumatic and non-traumatic film clips in the scanner. The authors found that increased activation of the inferior frontal gyrus during encoding predicted involuntary memories of highly emotional clips over the following week. Importantly, activation of the inferior frontal gyrus has also been implicated in voluntary recall of negative stimuli (Kensinger & Corkin, 2004), indicating that common mechanisms operate in both involuntary and voluntary retrieval. In a positron emission tomography (PET) study, Hall, Gjedde, and Kupers (2008) found differences in brain activation between involuntary and voluntary recall that were likely related to retrieval effort (e.g., increased activation of right prefrontal cortex in voluntary compared with voluntary retrieval), whereas areas related to retrieval success showed similar activation under both retrieval conditions. Recently, Hall et al. (in press) employed a paradigm similar to the one developed by Berntsen et al., (2013) in an fMRI study. During encoding, all participants heard sounds, half of which were paired with a picture of a scene while the other half were presented with no scene-pairing. Paired and unpaired sounds were presented again, while participants were in the scanner. In the voluntary condition, participants were asked to retrieve memories to all the sounds, while in the involuntary condition they were asked solve a sound location task, which was shown to trigger involuntary memories. The voluntary retrieval condition was associated with greater activity in dorsal frontal regions compared with the involuntary condition, presumably reflecting enhanced retrieval effort, whereas other structures typically engaged in retrieval, including the medial temporal, ventral occipitotemporal, and ventral parietal regions were similarly engaged by both types of memories.

These findings suggest that voluntary and involuntary memories are sampled from the same underlying structures, albeit through different retrieval mechanisms. This is counter to some clinical models proposing that involuntary retrieval preferentially accesses emotionally stressful and poorly processed material, stored separately from voluntarily accessible material (e.g., Ehlers & Clark, 2000; Brewin, 2013). In the meta-analysis of Experiments 1 to 4, we did find a 3-way interaction between emotional scenes, retrieval condition, and time. However, this interaction reflected a higher frequency of memories for emotional scenes in the *voluntary* condition after one week compared with the involuntary condition, and the two retrieval conditions followed the same overall pattern.

Future research should investigate voluntary and involuntary memory for personally relevant material in populations at-risk for PTSD (for example, people who previously experienced trauma). We propose that the paradigm presented here would be ideal for testing these hypotheses, since it allows the control of both encoding and retrieval. In addition, the paradigm could accommodate the inclusion of diaries during the retention interval and thereby integrate previous methods, such as the trauma analogue paradigm (Davies & Clark, 1998). As mentioned in the introduction, research has shown that intrusive memories are a core symptom in disorders other than PTSD (Brewin et al., 2010). For example, social anxiety has been associated with intrusive memories of stressful social encounters (Carleton, Peluso, Collimore, & Asmundson, 2011), while patients suffering from major depression have been found to involuntarily retrieve memories of interpersonal problems, illness, and death (Newby & Moulds, 2012). An important question is whether such intrusions of personally experienced negative events are different in terms of their underlying mechanisms compared with intrusions of traumatic events in PTSD. On the surface it may appear that events involving threat of death or bodily harm are more traumatizing than, for example, social defeat. However, this may not always be true (Carleton et al., 2011), and it does not rule out that the same memory mechanisms (e.g., enhanced consolidation of emotional events) could explain intrusive memories across disorders.

Another avenue for future research is a more detailed mapping of the time-course of forgetting and consolidation. When exactly do the neutral scenes begin to lose accessibility in memory relative to emotional scenes? Is sleep required for the consolidation of emotional memories, as some research has suggested (Wagner et al., 2001)? Can the process of consolidation of emotional memories be halted by behavioural interventions and if so, what would be the best timing for such interventions (see for instance Holmes, James, Kilford, & Deeprose, 2010)? We propose that the present paradigm would be suitable for investigating such important questions. Finally, the study of positive emotion could be pursued with the present paradigm. For example, by varying the valence and arousal of the stimuli, the relationship between individual differences and these dimensions of emotional experience could be further investigated (Kuppens, Tuerlinckx, Russel, & Barrett, 2013).

Some limitations should be taken into account when evaluating the present results. First, in the present experiments, the events were unlikely to have personal relevance for the participants, which could limit their applicability to real-life traumatic experiences. More specifically, real-world cues are likely to be more effective than the sound cues used in the present experiments. In the real world, cues will often have some meaningful relation to the event they are associated with, whereas the pairings in our experiments were arbitrary. This is likely to have increased the rate by which the effect of the unique cues diminished over time, relative to what would be the case in real life. Also the scenes presented to participants were unlikely to have strong personal relevance and both sounds and scenes were presented only briefly. Brewin (2013) argued that such material is unlikely to provoke strong avoidance reactions such as dissociation. It is important for future research to develop stimuli that vary in their personal relevance and intensity, when making arguments specifically related to posttraumatic stress.

Second, another potential limitation is the parallel task during involuntary retrieval. While this task was deliberately easy and simple, it is nonetheless a visuo-spatial task. When such tasks are employed during – or immediately following – encoding of emotional material, they have been shown to reduce the number of intrusions recorded over the following week (e.g., Holmes et al., 2010; Stuart, Holmes, & Brewin, 2006). It is presently not clear how a parallel task would affect intrusions during retrieval, but one possibility is that the visual search itself could reduce the emotional impact of memories (Kemps & Tiggemann, 2007). However, it is important to underscore that this potential factor would have worked against the effects observed in the present studies. Thus, without the parallel task, the effects of emotion might have been even more pronounced.

In conclusion, we found that involuntary retrieval of emotional scenes is governed by two factors asserting their influence at different points in time. Immediately after an event, cue distinctiveness affects memory performance, but after 24 hours, it is the emotional intensity of the scenes that becomes important. Voluntary retrieval shows the same pattern, but with increased retrieval time and higher frequencies of memories relative to available cues. We propose that the presented paradigm simulates intrusive memories of emotional events and can be modified and extended to investigate other aspects of trauma such as the importance of personal relevance, narrative coherence, intervention, and predisposing individual characteristics.

Acknowledgements

The authors are grateful to Tine Bennedsen Gehrt for assisting with the data collection and to Tine Blom Nielsen for assisting with interrater reliability. The present research was supported by a grant (DNRF93) from the Danish National Research Foundation (DNRF) to the Center on Autobiographical Memory Research. DNRF had no influence on planning, conducting or reporting of the present research.

References

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders (3rd ed.)*. Washington, DC: Author.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders (3rd ed., revised)*. Washington, DC: Author.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders (4th ed.)*. Washington, DC: Author.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders (4th ed., text rev.)*. Washington, DC: Author.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed., text rev.)*. Washington, DC: Author.

Ball, C. T., & Little, J. R. (2006). A comparison of involuntary memory retrievals. *Applied Cognitive Psychology, 20*, 1167–1179. doi:10.1002/acp.1264

Berntsen, D. (1996). Involuntary autobiographical memories. *Applied Cognitive Psychology*, *10*, 435– 454. doi:10.1002/(SICI)1099-0720(199610)10:5_435::AID-ACP408_3.0.CO;2-L

Berntsen, D. (1998). Voluntary and involuntary access to autobiographical memory. *Memory*, *6*, 113-141.

Berntsen, D. (2001). Involuntary memories of emotional events: Do memories of traumas and extremely happy events differ? *Applied Cognitive Psychology, 15*, S135–S158. doi:10.1002/acp.838

Berntsen, D. (2009). *Involuntary autobiographical memories: An introduction to the unbidden past*. Cambridge, England: Cambridge University Press. doi:10.1017/CBO9780511575921

Berntsen, D. (2010). The unbidden past: Involuntary autobiographical memories as a basic mode of remembering. *Current Directions in Psychological Science, 19*, 138–142. doi:10.1177/0963721410370301

Berntsen, D., & Hall, N. M. (2004). The episodic nature of involuntary autobiographical memories. *Memory & Cognition, 32,* 789–803. doi:10.3758/BF03195869

Berntsen, D., & Rubin, D. C. (2008). The reappearance hypothesis revisited: Recurrent involuntary memories after traumatic events and in everyday life. *Memory & Cognition, 36,* 449-460. doi:10.3758/MC.36.2.449

Berntsen, D. & Rubin, D.C. (2014). Involuntary memories and dissociative amnesia: Assessing key assumptions in posttraumatic stress disorder research. *Clinical Psychological Science, 2(2),* 174-186. doi:10.1177/2167702613496241.

Berntsen, D., Staugaard, S. R., & Sørensen, L. M. T. (2013). Why am I remembering this now? Predicting the occurrence of involuntary (spontaneous) episodic memories. *Journal of Experimental Psychology: General, 142,* 426–444. doi:10.1037/a0029128

Bjork, R. A. (2003). Interference and forgetting. In J. H. Byrne (Ed.), *Encyclopedia of learning and memory, 2nd ed.*, (pp. 268-273). New York: Macmillan Reference USA.

Bradley, B. P., & Baddeley, A. D. (1990). Emotional factors in forgetting. *Psychological Medicine, 20(2),* 351-355. doi:10.1017/S0033291700017669

Bourne, C., Mackay, C. E., & Holmes, E. A. (2013). The neural basis of flashback formation: the impact of viewing trauma. *Psychological Medicine, 43(7),* 1521-1532. doi:10.1017/S0033291712002358.

Brewin, C. R. (2014) Episodic memory, perceptual memory, and their interaction: Foundations for a theory of posttraumatic stress disorder. *Psychological Bulletin, 140(1)*, 69-97. doi: 10.1037/a0033722.

Brewin, C. R., Gregory, J. D., Lipton, M., & Burgess, N. (2010). Intrusive images and memories in psychological disorders: Characteristics, neural basis, and treatment implications. *Psychological Review, 117*, 210–232. doi:10.1037/a0018113

Cahill, L., & McGaugh, J. L. (1998). Mechanisms of emotional arousal and lasting declarative memory. *Trends in Neurosciences, 21(7)*, 294-299.

Carleton, R. N., Peluso, D. L., Collimore, K. C., & Asmundson, G. J. (2011). Social anxiety and posttraumatic stress symptoms: the impact of distressing social events. *Journal of Anxiety Disorders, 25(1),* 49-57. doi:10.1016/j.janxdis.2010.08.002.

Conway, M. A. (2005). Memory and the self. *Journal of Memory and Language, 53,* 594-628. doi:10.1016/j.jml.2005.08.005.

Crick, F., & Mitchison, G. (1983). The function of dream sleep. *Nature, 304(5922),* 111–114. doi:10.1038/304111a0.

Davies, M. I., & Clark, D.M. (1998). Predictors of analogue post-traumatic intrusive cognitions. *Behavioural and Cognitive Psychotherapy, 26,* 303–314.

Dudai, Y. (2004). The neurobiology of consolidations, or, how stable is the engram? *Annual Review of Psychology, 55,* 51-86. doi:10.1146/annurev.psych.55.090902.142050

Ehlers, A., & Clark, D. M. (2000). A cognitive model of posttraumatic stress disorder. *Behaviour Research and Therapy, 38,* 319–345. doi:10.1016/S0005-7967(99)00123-0

Ehlers, A., Hackmann, A., & Michael, T. (2004). Intrusive re-experiencing in post-traumatic stress disorder: Phenomenology, theory, and therapy. *Memory, 12*, 403–415. doi:10.1080/09658210444000025

Ehlers, A., Mauchnik, J., & Handley, R. (2012). Reducing unwanted trauma memories by imaginal exposure or autobiographical memory elaboration: An analogue study of memory processes. *Journal of Behavior Therapy and Experimental Psychiatry, 43*, S67–S75. doi:10.1016/j.jbtep.2010.12.009

Ferré, P. (2002). Advantage for emotional words in immediate and delayed memory tasks: Could it be explained in terms of processing capacity? *The Spanish Journal of Psychology*, *5(2)*, 78-89. doi:

Galton, F. (1907). *Inquiries into human faculty and its development*. London: J. M. Dent and Sons.

Hall, N. M., Gjeddes, A., & Kupers, R. (2008). Neural mechanisms of voluntary and involuntary recall: a PET study. *Behavioral Brain Research, 186(2)*, 261-272. doi: 10.1016/j.bbr.2007.08.026.

Hall, S., A., Rubin, D. C., Miles, A., Davis, S. W., Wing, E. A., Cabeza, R., & Berntsen, D. (in press). The neural basis of involuntary episodic memories. *Journal of Cognitive Neuroscience*.

Holmes, E. A., Brewin, C. R., & Hennessy, R. G. (2004). Trauma films, information processing, and intrusive memory development. *Journal of Experimental Psychology: General*, *133*, 3–22. doi:10.1037/0096-3445.133.1.3

Holmes, E. A., James, E. L., Kilford, E. J., & Deeprose, C. (2010). Key steps in developing a cognitive vaccine against traumatic flashbacks: Visuospatial Tetris versus verbal Pub Quiz. *PLoS ONE, 5*, e13706. doi:10.1371/journal.pone.0013706

Hunt, R.R. & Worthen, J.B. (Eds.) (2006). *Distinctiveness and memory*. New York: Oxford University Press.

Kebeck, G., & Lohaus, A. (1986). Effect of emotional arousal on free recall of complex material. *Perceptual and Motor Skills, 63(2),* 461-462.

Kemps, E., & Tiggemann, M. (2007). Reducing the vividness and emotional impact of distressing autobiographical memories: The importance of modality-specific interference. *Memory*, *15*, 412-422. doi:10.1080/09658210701262017

Kensinger, E.A., & Corkin, S. (2004). Two routes to emotional memory: Distinct neural processes for valence and arousal. *Proceedings of the National Academy of Sciences, 101*, 3310-3115. doi:10.1073pnas.0306408101.

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

Kleinsmith, L.J., & Kaplan, S. (1964). The interaction of arousal and recall interval in nonsense syllable paired-associate learning. *Journal of Experimental Psychology, 67*, 124-126. doi: 10.1037/h0045203. doi:10.1037/h0040288.

Kleinsmith, L. J., Kaplan, S., & Tarte, R. D. (1963). The relationship of arousal to short- and long-term verbal recall. *Canadian Journal of Psychology, 17*, 393–397.

Kuppens, P., Tuerlinckx, F., Russell, J. A., & Barrett, L. F. (2013). The relation between valence and arousal in subjective experience. *Psychological Bulletin, 139*, 917-940. doi:10.1037/a0030811.

Kvavilashvili, L., & Mandler, G. (2004). Out of one’s mind: A study of involuntary semantic memories. *Cognitive Psychology*, *48*, 47–94. doi:10.1016/S0010-0285(03)00115-4

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.

LaBar, K. S., & Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nature Reviews Neuroscience, 7*, 54-64. doi:10.1038/nrn1825.

Levinger, G., & Clark, J. (1961). Emotional factors in the forgetting of word associations. *Journal of Abnormal Psychology, 62*, 99-105.

Levonian, E. (1967). Retention of Information in Relation to Arousal during Continuously-Presented Material. *American Educational Research Journal, 4(2),* 103-116.

Mace, J. H. (2004). Involuntary autobiographical memories are highly dependent on abstract cueing: The Proustian view is incorrect. *Applied Cognitive Psychology, 18*, 893–899. doi:10.1002/acp.1020

Mace, J. H. (Ed.). (2007). *Involuntary memory*. Malden. MA: Blackwell.

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

McGaugh, J. L. (2003). *Memory and emotion: The making of lasting memories*. New York: Colombia University Press.

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

Moscovitch, M. (1995). Recovered consciousness: A hypothesis concerning modularity and episodic memory. *Journal of Clinical and Experimental Neuropsychology, 17(2),* 276-90. doi:10.1080/01688639508405123

Newby, J. M., & Moulds, M. L. (2012). A comparison of the content, themes and features of intrusive memories and rumination in major depressive disorder. *British Journal of Clinical Psychology, 51(2),* 197-205. doi:10.1111/j.2044-8260.2011.02020.x

Parkin, A. J., Lewinson, J., & Folkard, S. (1982). The influence of emotion on immediate and delayed retention: Levinger & Clark reconsidered. *British Journal of Psychology, 73(3),* 389-393. doi:10.1111/j.2044-8295.1982.tb01821.x

Pearson, D. G., Ross, F. D. C., & Webster, V. L. (2012). The importance of context: Evidence that contextual representations increase intrusive memories. *Journal of Behavior Therapy and Experimental Psychiatry, 43,* 573–580. doi:10.1016/j.jbtep.2011.07.009

Porter, S., & Peace, K. (2007). The scars of memory: A prospective, longitudinal investigation of the consistency of traumatic and positive emotional memories in adulthood. *Psychological Science, 18,* 435-441. doi:10.1111/j.1467-9280.2007.01918.x

Roediger, H. L., & Guynn, M. J. (1996). Retrieval processes. In E. L. Bjork & R. A. Bjork (Eds.), *Human memory* (197-236). San Diego: Academic Press.

Roediger, H. L., Weinstein, Y., & Agarwal, P. K. (2010). Forgetting: Preliminary considerations. In S. Della Sala (Ed.), *Forgetting* (pp. 1-22). Hove, UK: Psychology Press.

Rubin, D. C. (1995). *Memory in oral traditions. The cognitive psychology of epic, ballads, and counting-out rhymes*. New York, NY: Oxford University Press.

Salomons, T. V., Osterman, J. E., Gagliese, L., Katz, J. (2004). Pain flashbacks in posttraumatic stress disorder. *The Clinical Journal of Pain, 20(2),* 83-7.

Schaefer, A., & Philippot, P. (2005). Selective effects of emotion on the phenomenal characteristics of autobiographical memories. *Memory, 13(2),* 148-160. doi:10.1080/09658210344000648

Schlagman, S., & Kvavilashvili, L. (2008). Involuntary autobiographical memories in and outside the laboratory: How different are they from voluntary autobiographical memories? *Memory and Cognition, 36(5),* 920-932.

Sharot, T. & Yonelinas, A.P. (2008). Differential time-dependent effects of emotion on recollective experience and memory for contextual information. *Cognition, 106(1)*, 538–547. doi:10.1016/j.cognition.2007.03.002

Sierra, M., & Berrios, G. E. (1999). Towards a neuropsychiatry of conversive hysteria. *Cognitive Neuropsychiatry, 4*, 267-287.

Singer, J. L. (1966). *Daydreaming. An introduction to the experimental study of inner experience*. New York: Random House.

Stuart, A. D. P., Holmes, E. A., & Brewin, C. R. (2006). The influence of a visuospatial grounding task on intrusive images of a traumatic film. *Behaviour Research and Therapy, 44*, 611-619. doi:10.1016/j.brat.2005.04.004

Steel, C., Holmes, E. A. (2007). The role of involuntary memories in posttraumatic disorder and psychosis. In J. H. Mace (ed.) *Involuntary memory*, 68-86. Malden: Blackwell Publishing.

Stickgold, R., Hobson, J. A., Fosse, R., & Fosse, M. (2001). Sleep, learning, and dreams: Off-line memory reprocessing. *Science, 294(5544)*, 1052-1057. doi:10.1126/science.1063530

St. Jacques, P.L., & Levine, B. (2007). Ageing and autobiographical memory for emotional and neutral events. *Memory, 15*, 129-144. doi:10.1080/09658210601119762

Talarico, J. M., LaBar, K. S., & Rubin, D. C. (2004). Emotional intensity predicts autobiographical memory experience. *Memory & Cognition, 32*, 1118–1132. doi:10.3758/BF03196886

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

Tulving, E. (1974). Cue-dependent forgetting. *American Scientist, 62(1)*, 74-82.

Tulving, E., & Thompson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80(5),* 352-373.

Underwood, B. J. (1957). Interference and forgetting. *Psychological Review, 64(1),* 49–60. doi:10.1037/h0044616

Underwood, B.J. and Postman, L. (1960). Extra-experimental sources of interference in forgetting, *Psychological Review, 67(2),* 73-95. doi:10.1037/h0041865

Wagner, U., Gais, S., & Born, J. (2001). Emotional memory formation is enhanced across sleep intervals with high amounts of rapid eye movement sleep. *Learning and Memory, 8(2),* 112-119.

Watkins, O. C. and Watkins, M. J. 1975. Buildup of proactive inhibition as a cue overload effect. *Journal of Experimental Psychology: Human Learning and Memory, 1,* 442–52.

Wixted, J. T. (2004). The psychology and neuroscience of forgetting. *Annual Review of Psychology, 55, 235-269*. doi:10.1146/annurev.psych.55.090902.141555

Table 1. Means, standard deviations, and *F*-statistics for the subjective characteristics of retrieved scenes in Experiment 1[1] and 2[1].

	Involuntary		Voluntary		Main effects		Interaction
	Emo (*SD*)	Neu (*SD*)	Emo (*SD*)	Neu (*SD*)	INV vs. VOL	Emo vs. Neu	
Experiment 1	$N = 12$		$N = 16$		$F(1,26)$	$F(1,26)$	$F(1,26)$
Intensity	4.37 (0.43)	1.45 (0.40)	4.12 (0.61)	1.35 (0.37)	1.59	590.30**	0.44
Bodily reactions	2.82 (0.98)	1.19 (0.20)	2.77 (1.22)	1.11 (0.17)	0.08	63.05**	0.01
Specificity	3.73 (0.91)	4.52 (0.59)	3.68 (0.64)	4.54 (0.37)	0.01	42.89**	0.07
Reaction time	3834 (1296)	3549 (1044)	6969 (3460)	5519 (1920)	14.35*	2.59	1.17
Experiment 2	$N = 15$		$N = 16$		$F(1,29)$	$F(1,29)$	$F(1,29)$
Intensity	4.18 (0.72)	1.63 (0.55)	3.67 (1.23)	1.43 (0.32)	2.36	218.97**	0.89
Bodily reactions	2.90 (1.30)	1.21 (0.24)	2.69 (1.13)	1.11 (0.19)	0.41	64.43**	0.07
Specificity	3.42 (0.96)	4.19 (0.71)	3.83 (0.84)	4.48 (0.55)	2.57	17.43**	0.13
Reaction time	3812 (1338)	3193 (1083)	9305 (4513)	7036 (4671)	24.71**	3.36	1.10

The interaction is between valence of the scenes and retrieval condition. INV = involuntary retrieval condition, VOL = voluntary retrieval condition, Emo = emotional scenes, Neu = neutral scenes. * $p = .001$, ** $p < .0001$. [1]Based on unique cue trials only.

Table 2. Means and standard deviations for indiscriminately cued memories.

		Involuntary		Voluntary	
		Total M (*SD*)	Proportion (%)	Total M (*SD*)	Proportion (%)
Experiment 1					
	UniCue_EmoScene	1.25 (1.48)	18.86	1.69 (1.66)	25.73
	UniCue_NeuScene	0.50 (0.89)	11.56	0.88 (0.81)	17.37
Experiment 2					
	UniCue_EmoScene	0.75 (1.34)	13.17	1.06 (1.53)	20.63
	UniCue_NeuScene	0.94 (0.85)	17.07	1.19 (1.60)	16.69
Experiment 3					
	UniCue_EmoScene	2.00 (1.67)	40.96	2.31 (1.74)	46.46
	UniCue_NeuScene	1.19 (1.52)	24.88	1.50 (1.51)	34.27
Experiment 4					
	UniCue_EmoScene	1.75 (1.57)	53.96	2.5 (1.83)	50.24
	UniCue_NeuScene	0.75 (1.44)	23.44	1.13 (1.02)	48.35

UniCue_EmoScene = unique cue and emotional scene, UniCue_NeuScene = unique cue and neutral scene.

Table 3. Means, standard deviations, and *F*-statistics for the subjective characteristics of retrieved scenes in Experiment 3[1] and 4[2].

	Involuntary		Voluntary		Main effects		Interaction
	Emo (*SD*)	Neu (*SD*)	Emo (*SD*)	Neu (*SD*)	INV vs. VOL	Emo vs. Neu	
Experiment 3	*N* = 15		*N* = 14		*F*(1,27)	*F*(1,27)	*F*(1,27)
Intensity	3.66 (0.42)	1.58 (0.42)	3.97 (0.70)	1.20 (0.24)	0.09	397.76***	8.24*
Bodily reactions	2.54 (0.83)	1.20 (0.34)	2.81 (0.83)	1.19 (0.24)	0.58	90.00***	0.78
Specificity	3.71 (0.88)	4.32 (0.52)	3.35 (1.03)	4.25 (0.61)	0.67	32.29***	1.19
Reaction time	3774 (1657)	2938 (1223)	8533 (4622)	7638 (3345)	29.36***	1.70	0.00
Experiment 4	*N* = 11		*N* = 13		*F*(1,22)	*F*(1,22)	*F*(1,22)
Intensity	3.45 (0.72)	1.43 (0.63)	4.06 (0.86)	1.56 (0.51)	3.03	149.18***	1.64
Bodily reactions	2.46 (0.94)	1.17 (0.26)	2.43 (0.84)	1.12 (0.19)	0.05	62.12***	0.00
Specificity	3.52 (0.91)	2.98 (1.19)	2.97 (1.08)	3.33 (0.98)	0.11	0.10	2.46
Reaction time	3902 (1670)	4542 (2736)	7941 (2854)	6636 (2017)	14.85**	0.35	2.98

The interaction is between valence of the scenes and retrieval condition. INV = involuntary retrieval condition, VOL = voluntary retrieval condition, Emo = emotional scenes, Neu = neutral scenes. *p = .01, ** p = .001, *** p < .0001. [1]Based on unique cue trials only. [2]Based on repeated cue trials only.

Appendix

Means and standard deviations (SD) for the frequency of emotional and neutral memories across all four experiments.

	Involuntary				Voluntary			
	UNI_EMO	UNI_NEU	REP_EMO	REP_NEU	UNI_EMO	UNI_NEU	REP_EMO	REP_NEU
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Experiment 1	5.44 (4.91)	5.31 (4.06)	3.00 (3.63)	2.44 (3.41)	5.94 (1.88)	6.19 (1.60)	4.13 (2.22)	2.44 (3.41)
Experiment 2	3.81 (2.43)	5.00 (2.42)	3.44 (4.80)	2.06 (2.27)	4.31 (2.02)	6.25 (1.57)	3.44 (2.37)	3.44 (2.34)
Experiment 3	4.88 (2.25)	4.06 (2.86)	4.25 (3.92)	1.94 (2.05)	4.88 (2.28)	4.0 (2.22)	6.81 (3.58)	3.44 (2.39)
Experiment 4	2.56 (1.79)	1.25 (1.69)	3.50 (3.10)	2.38 (2.90)	4.63 (2.09)	2.44 (1.97)	5.50 (4.29)	2.06 (1.39)

UNI_EMO = unique cue, emotional scene, UNI_NEU = unique cue, neutral scene, REP_EMO = repeated cue, emotional scene, REP_NEU = repeated cue, neutral scene

Figure 1.

Frequency of scenes recorded relative to the number of cues for each trial type in Experiment 1.

Error bars: 1 standard error.

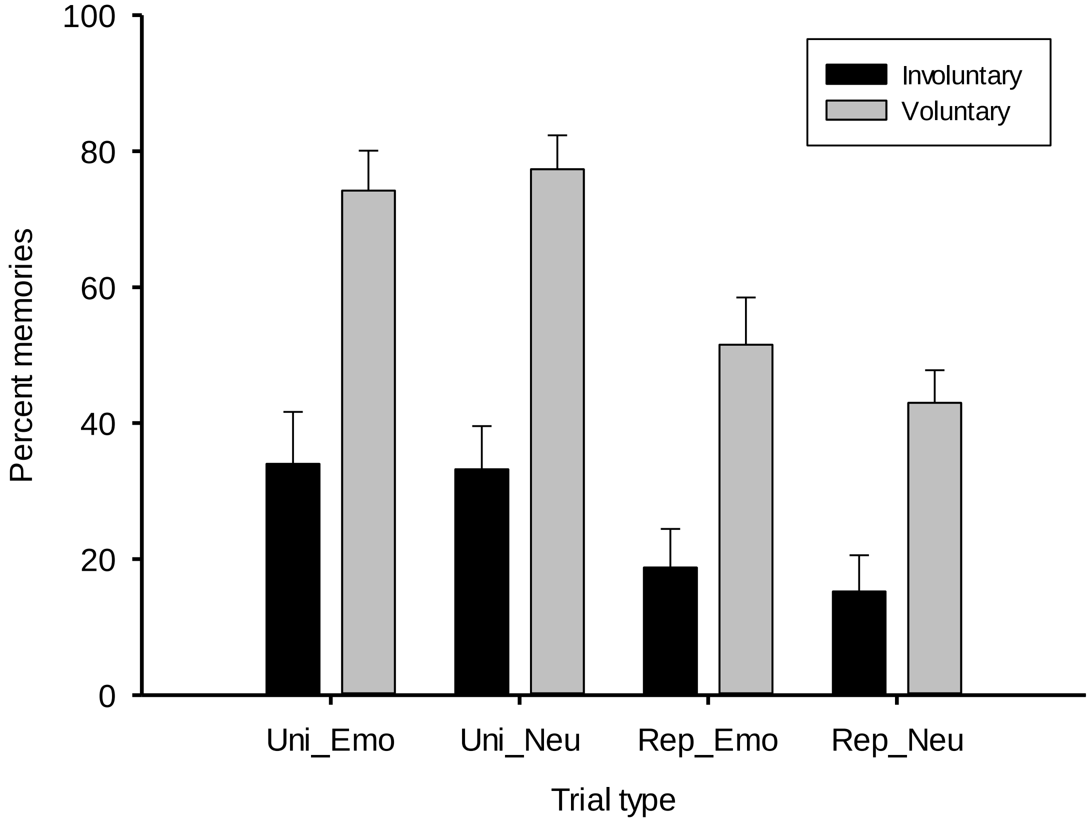


Figure 2.

Frequency of scenes recorded relative to the number of cues for each trial type in Experiment 2.

Error bars: 1 standard error.

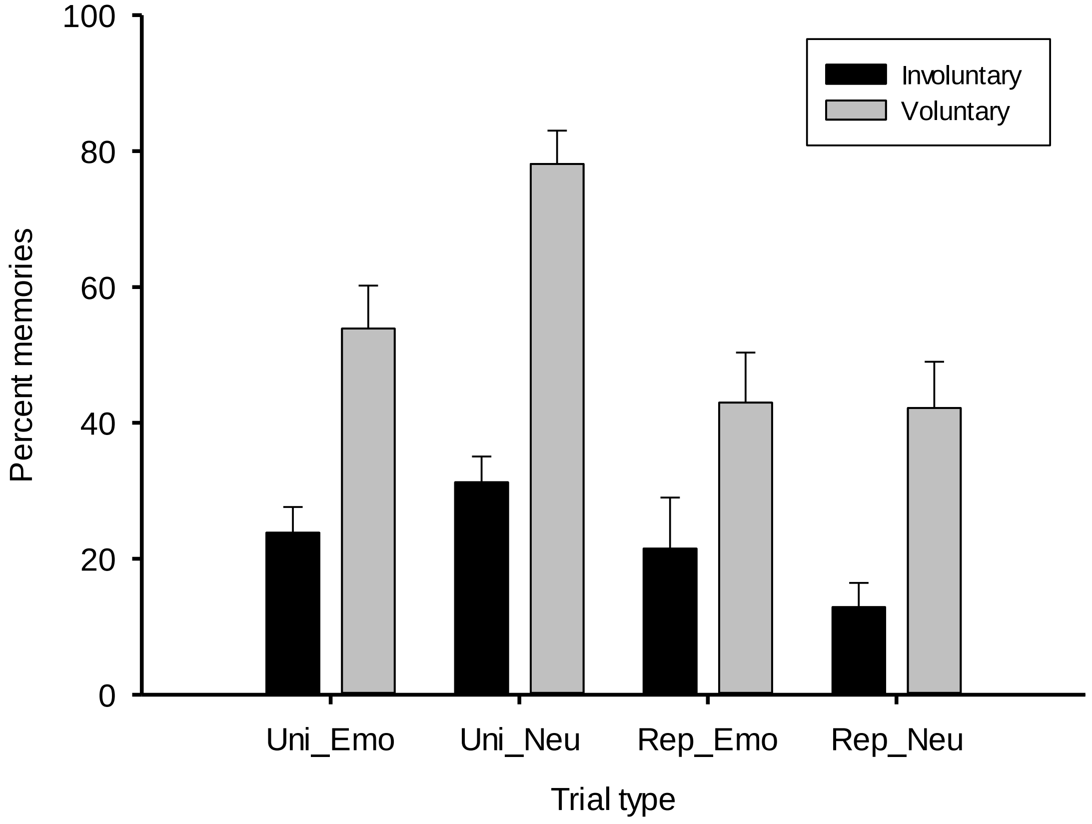


Figure 3.

Frequency of scenes recorded relative to the number of cues for each trial type in Experiment 3.

Error bars: 1 standard error.

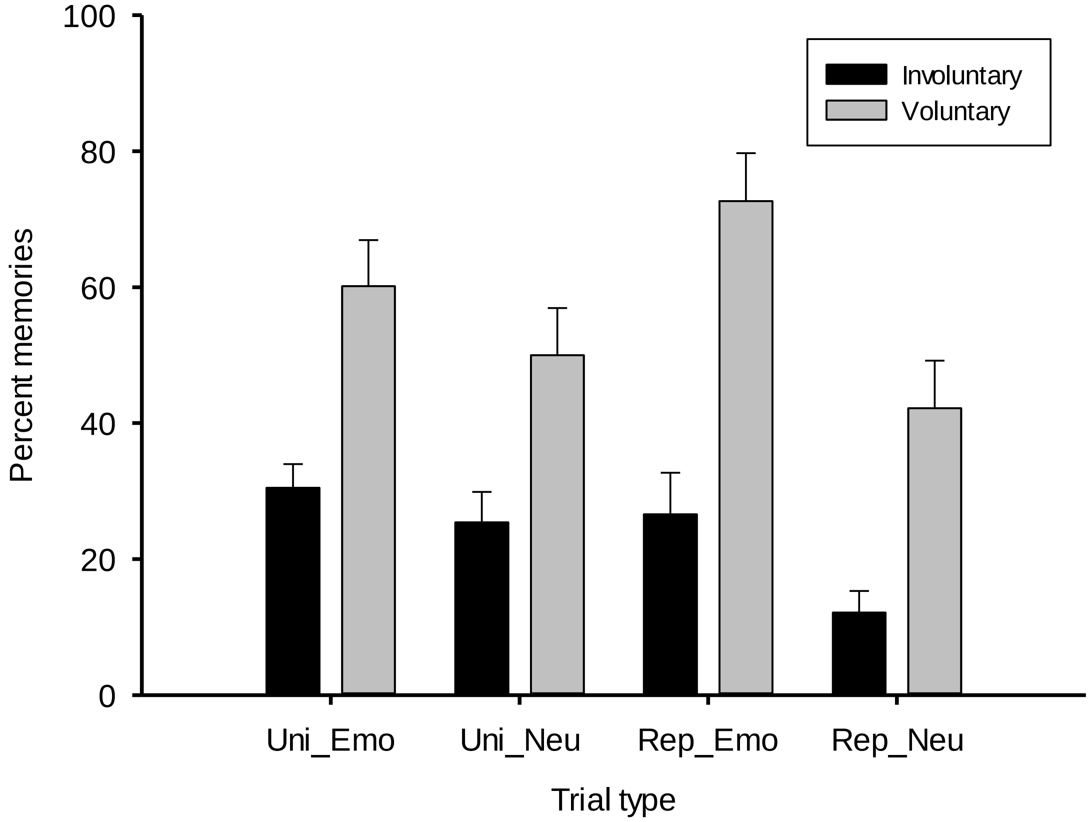


Figure 4.

Frequency of scenes recorded relative to the number of cues for each trial type in Experiment 4.

Error bars: 1 standard error.

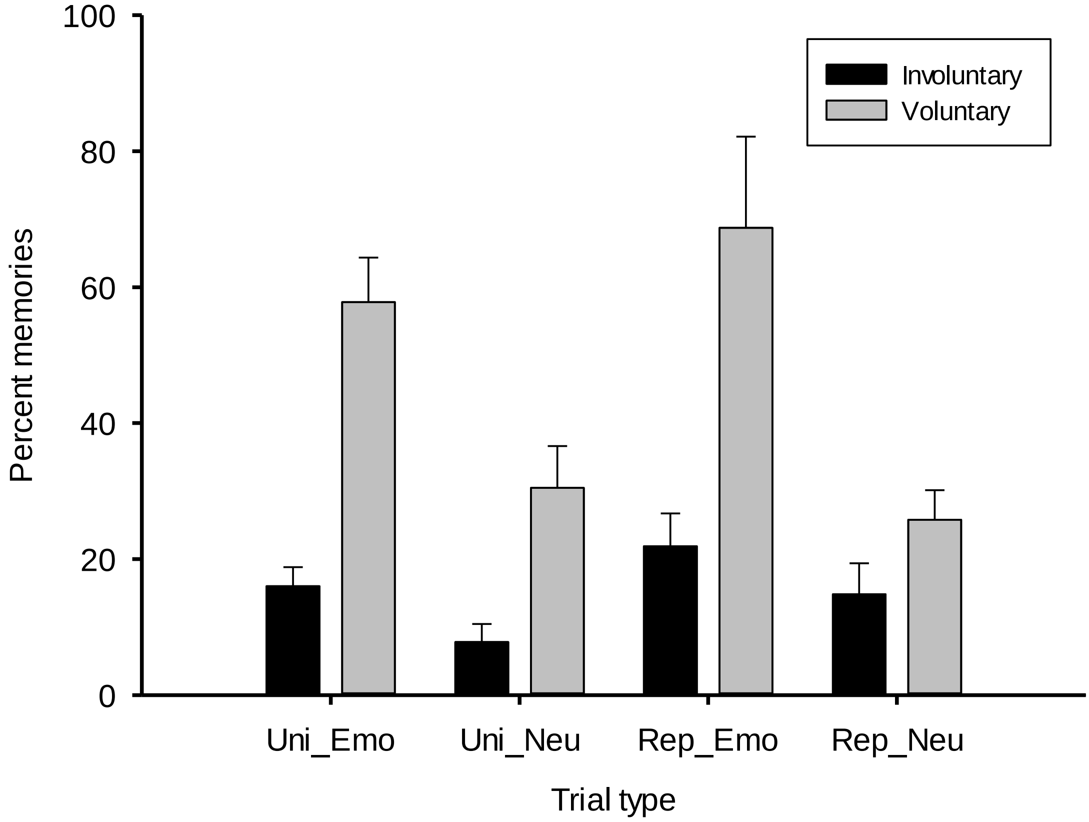
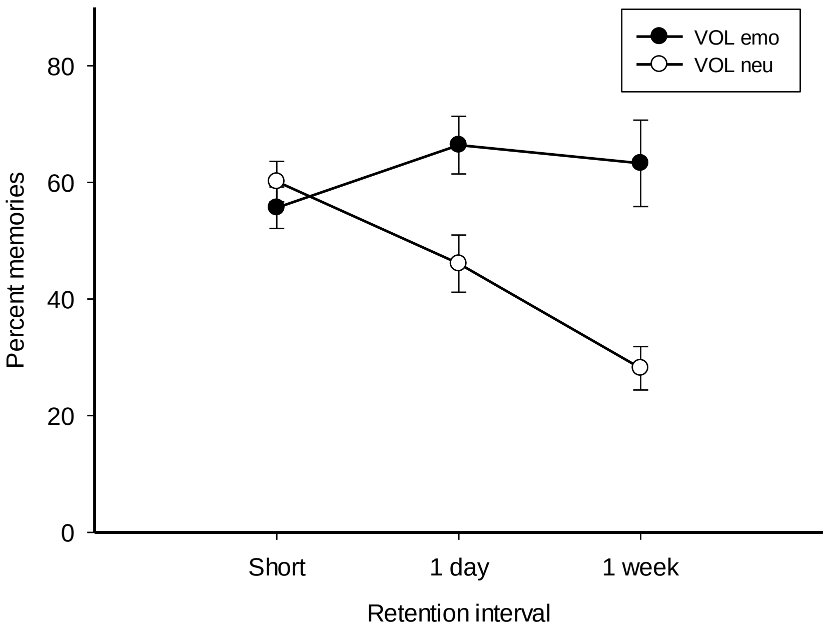


Figure 5.

Frequency of emotional and neutral scenes recorded relative to the number of cues across all participants and all time points. Top panel: Voluntary condition, bottom panel: involuntary condition. Error bars: 1 standard error.



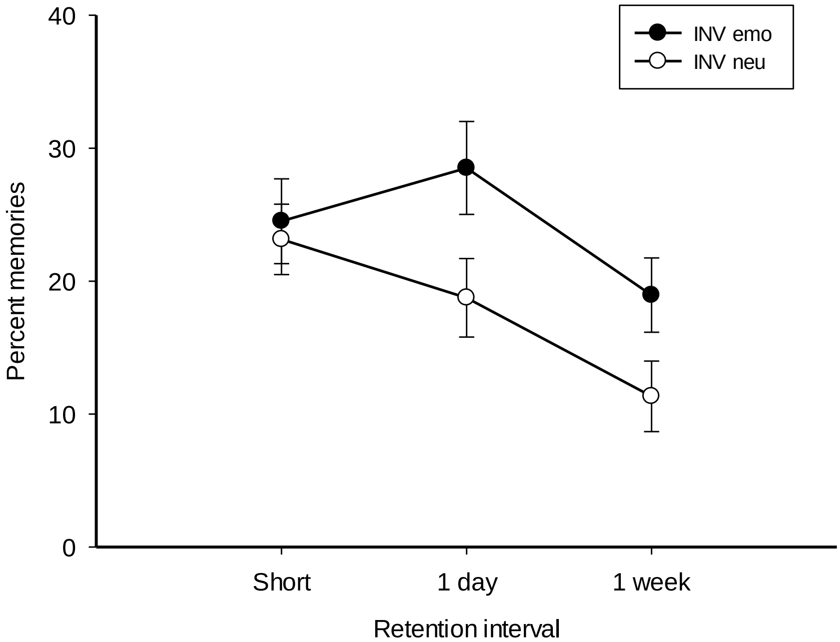


Figure 6.

Frequency of uniquely cued and repeatedly cued scenes recorded relative to the number of cues across all participants and all time points. Top panel: Voluntary condition, bottom panel: involuntary condition. Error bars: 1 standard error.