



AARHUS UNIVERSITY



Coversheet

This is the accepted manuscript (post-print version) of the article.

Contentwise, the accepted manuscript version is identical to the final published version, but there may be differences in typography and layout.

How to cite this publication

Please cite the final published version:

Staugaard, S. R., & Berntsen, D. (2019). Retrieval intentionality and forgetting: How retention time and cue distinctiveness affect involuntary and voluntary retrieval of episodic memories. *Memory & Cognition*. <https://doi.org/10.3758/s13421-019-00904-w>

Publication metadata

Title:	Retrieval intentionality and forgetting: How retention time and cue distinctiveness affect involuntary and voluntary retrieval of episodic memories
Author(s):	Staugaard, Søren Risløv; Berntsen, Dorthe.
Journal:	Memory & Cognition
DOI/Link:	https://doi.org/10.3758/s13421-019-00904-w
Document version:	Accepted manuscript (post-print)

General Rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

If the document is published under a Creative Commons license, this applies instead of the general rights.

Retrieval Intentionality and Forgetting: How Retention Time and Cue Distinctiveness

Affect Involuntary and Voluntary Retrieval of Episodic Memories

Søren Risløv Staugaard, Dorthe Berntsen

Center on Autobiographical Memory Research, Department of Psychology and Behavioral

Sciences, Aarhus University

Corresponding author:

Søren R. 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

Forgetting, understood as a measurable reduction in memory accessibility over time, has been studied extensively in episodic memory for more than 130 years. However, this research has typically focused on voluntary memory – that is, information retrieved intentionally. Few studies have examined forgetting in relation to involuntary memories, that is, memories coming to mind spontaneously with no preceding attempt of retrieval. The purpose of the present studies was to investigate the effects of cue distinctiveness and the passage of time on the accessibility of involuntary and voluntary memories for pictures of scenes. For both types of retrieval, we examined the frequency of correct memories after a few minutes, 1 day, and 7 days; in Study 2, also after 3 days. Across both studies and both types of retrieval, distinct cues yielded better memory access than indistinct cues, and memory frequency dropped systematically with increased retention time. At the shortest retention interval, voluntary retrieval led to greater memory access than involuntary retrieval, but after one week, this advantage had disappeared. The findings suggest that memories for events become more cue-dependent over time, which limits the beneficial effects of strategic search associated with voluntary retrieval at longer delays.

Keywords: Episodic memory, forgetting, involuntary memories, autobiographical memory, retrieval intentionality

Retrieval Intentionality and Forgetting: How Cue Distinctiveness and Retention Time Affects Involuntary and Voluntary Retrieval of Episodic Memories

In his seminal book, Herman Ebbinghaus (1885) introduced experimental research on forgetting and discovered the forgetting curve, showing a rapid loss of information at short delays after which the forgetting rate becomes slower. This curve has been replicated numerous times for a range of material, including naturalistic events (Rubin & Wenzel, 1996). Although less frequently noted, Ebbinghaus (1885) used a method (the savings method) that was purposely indifferent to the form of retrieval (e.g., whether intentional or unintentional). In contrast, most research on forgetting that has been conducted since then has studied forgetting through various forms of intentional, goal-directed retrieval (e.g., Rubin & Wenzel, 1996, for a review) or unintentional and also unconscious memory (that is, implicit memory e.g., McBride & Doshier, 1997; McBride, Doshier, & Gage, 2001). Here we study the forgetting of conscious memories of scenes that are retrieved intentionally versus unintentionally. We use the term forgetting to indicate a reduction in memory accessibility (but not necessarily reduced availability, in terms of lost traces, Tulving & Pearlstone, 1966) as a function of time.

In the field of autobiographical memory, a distinction is often made between voluntary and involuntary retrieval, where the key difference is whether a person intended to remember or not (e.g., Berntsen, 1996, 2009, 2010; Mace, 2007; Schlagman & Kvavilashvili, 2008). Contrary to their voluntary counterparts, involuntary memories enter awareness unexpectedly and without any conscious effort at retrieval (Berntsen & Hall, 2004; Johannessen & Berntsen, 2010; for a review, see Berntsen, 2009). In spite of these differences concerning retrieval, there are many similarities between involuntary and voluntary memories (Berntsen, 2010; Mace, Atkinson, Moeckel, & Torres, 2011). One

important similarity appears to be the rate of forgetting over time. When studied in naturalistic settings, the frequency of both types of memory drops with increase in retention time, and with distributions that can be fitted by a standard retention function, accounting for more than 90% of the variance (Berntsen, 1998; Schlagman, Kliegel, Schulz, & Kvavilashvili, 2009), consistent with the classical forgetting curve (Ebbinghaus, 1885). Compared with voluntary memories, the forgetting of involuntary memories tends to be a little faster over time, yielding a slightly steeper slope of the forgetting curve (Berntsen, 1998, 2009). Also, while participants assign higher confidence ratings to their distant involuntary memories, the accuracy of involuntary and voluntary memory is very similar (Mace et al., 2011).

These earlier studies used a diary method, which has clear strengths but also some notable weaknesses. Firstly, the method ensures high ecological validity, but the lack of control over the encoding of personal events can be a source of bias especially in relation to forgetting. For example, it is unclear whether the contents of the events retrieved involuntarily and voluntarily are fully comparable. Secondly, the diary method allows little control over retrieval, such as cueing. For example, in the three studies reviewed above, voluntary memories were retrieved in response to word cues selected by the experimenter during a single session, while involuntary memories were retrieved in response to environmental cues during an extended period of daily living (Berntsen, 1998; Mace et al., 2011; Schlagman et al., 2009). Since cue characteristics have been shown to be a determining factor in retrieval success (e.g., Berntsen, Staugaard, & Sørensen, 2013; Hunt & Smith, 1996; Mäntylä & Nilsson, 1988), experimental control over cueing is important when investigating forgetting.

The goal of the present studies is to examine the forgetting rate of voluntary and involuntary memories using an experimental paradigm that controls both the encoding and

retrieval phase and provides similar cueing conditions for involuntary and voluntary memories. One important reason to control cueing is that involuntary and voluntary retrieval may show a different pattern of retrieval depending on the distinctiveness of the cue. We examined this in a recent experiment, where we manipulated cue distinctiveness by pairing repeated and unique sound cues with repeated and unique picture scenes (Berntsen et al., 2013). During retrieval, participants assigned to either an involuntary or a voluntary retrieval condition recorded equally high frequencies of memories in response to unique cues. However, participants in the voluntary condition also recorded a high frequency of memories in response to repeated cues, while participants in the involuntary condition recorded a very low frequency of memories to repeated cues. We interpreted this to mean that the strategic search process underlying voluntary memory is able to access more general (non-specific) information, while involuntary retrieval depends more on associative strength, favoring highly distinct cues (Berntsen et al., 2013).

Taken together, there is evidence from naturalistic studies that voluntary and involuntary retrieval is equally accurate (Mace et al., 2011) and show a similar drop in accessibility over time (Berntsen, 1998, 2009; Schlagman et al., 2009). However, experimental evidence (Berntsen et al., 2013; Staugaard & Berntsen, 2014) indicates that retrieval strategy interacts with cue distinctiveness. Specifically, voluntary retrieval leads to more memories than involuntary retrieval when cues are rendered indistinct. This advantage could be due to an increased capacity for interference resolution associated with voluntary retrieval. Interference resolution refers to a process of inhibiting competing information, when trying to remember specific information (Wimber, Rutschmann, Greenlee, & Bäuml, 2009). For example, when trying to remember your last visit to the dentist, details about other dental visits might interfere with your memory unless they are successfully inhibited. In the

present study, we wanted to examine how this strategic advantage affects the frequency of memories as a function of the passage of time.

To the best of our knowledge, no prior study has measured the frequency of involuntary and voluntary memory as a function of both cue distinctiveness and the passage of time, using an experimental approach. We recently investigated the effects of cue distinctiveness and retention interval on voluntary and involuntary memory for emotional and neutral scenes (Staugaard & Berntsen, 2014). However, the primary focus of Staugaard and Berntsen (2014) was emotion and not forgetting rate. For this reason, the two retrieval conditions were not completely balanced, and we did not analyze the effects of cueing and retrieval condition on reduced access over time independently of emotion.

The aim of the present studies is to examine the forgetting rate of involuntary and voluntary memories, while experimentally controlling for characteristics of the to-be-remembered material as well as cueing conditions. We first reanalyze data from Staugaard and Berntsen (2014) with a special emphasis on the possible interaction between cue distinctiveness and retrieval condition on memory accessibility over time (i.e., forgetting) while focusing on the neutral memories. We next conduct a new study, specifically tailored to investigate how cue distinctiveness and retention time affect involuntary and voluntary memories. Based on our prior studies and the concept of interference resolution, we expect voluntary memory to show greater retrieval success in response to indistinctive cues. However, whether this effect is influenced by the passage of time has not been studied before, and is an open question for the present study to explore.

The present studies

The present studies were based on a paradigm developed specifically to investigate the effects of cueing on voluntary and involuntary retrieval (Berntsen, Staugaard & Sørensen, 2013; Staugaard & Berntsen, 2014). While the specific stimuli and parts of the procedure

were different in Study 1 and Study 2 (to be described below), several aspects were common to both studies: Participants were randomly assigned to two retrieval conditions (voluntary vs. involuntary) and to separate retention intervals, ranging in duration from a few minutes to one week. Participants completed three experimental phases, consisting of encoding, retrieval, and recognition, in this order. In the encoding phase, the participants were shown photographs of scenes paired with sounds. The sounds would later be used as cues for the memories of the scenes in the retrieval phase. Importantly, the sound cues differed in their distinctiveness. During retrieval, participants in the involuntary condition performed a simple and repetitive visual attention task, while listening to the sound cues from encoding. They were instructed to pause the attention task if they spontaneously experienced an image appear in their minds. Participants in the voluntary condition simply listened to the sound cues and tried to remember the associated scene.

Study 1

Method

Study 1 is based on an analysis of data from Staugaard and Berntsen (2014) that has not been previously reported. The full methods, materials, and design of this prior study can be found in Staugaard and Berntsen (2014). One hundred twenty-eight participants were randomly assigned to one of two retrieval conditions (voluntary vs. involuntary) and to one of three retention intervals (immediate vs. 24-hours vs. 1-week). Since the instructions were deliberately vague in the involuntary condition to obscure the true purpose of the study, some participants would misunderstand them and report images that were uncodable. Two participants were excluded for not providing any codable memories. Participants completed a separate encoding, retrieval, and recognition phase, described below.

Materials. Pictures were taken from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2008) and their selection was based on the ratings

included with the set. Negative scenes had a mean valence rating of 1.81 ($SD = 0.09$) and mean arousal rating of 6.17 ($SD = 0.42$), while neutral scenes had a mean valence rating of 4.57 ($SD = 0.11$) and a mean arousal rating of 3.27 ($SD = 0.48$). Examples of negative scenes included scarred or mutilated bodies and children starving. Neutral scenes included people in neutral settings as well as everyday objects. In addition, 32 pictures of scenes were taken from various sources on the internet 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). All sounds were normalized and cut short at four seconds. While 16 of the sounds were dissimilar (e.g., a person coughing or a telephone ringing), the remaining 32 sounds consisted of 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. To ensure this was the case, a panel of independent judges ($N = 9$) rated each sound on a scale of 1-5, with 1 = highly negative, 3 = neutral and 5 = highly positive. The ratings of the independent judges were consistent (intraclass correlation = .796) and confirmed the neutral quality of the sounds ($M = 2.89$, $SD = 0.27$). Finally, an additional 32 sounds were collected using the same procedure described above, to be used as filler sounds in the involuntary condition.

Encoding. Participants watched 32 trials consisting of 16 emotional and 16 neutral scenes paired with the 16 unique and 16 of the repeated sound cues consisting of two sets of similar sounds, for example, repeated sounds of birds singing and dogs barking. Each trial was preceded by a central fixation cross for 2 seconds. Then the picture-sound pair was displayed for 4 seconds. Pictures were presented at half their original size. Participants were

then asked to indicate on a 5-point scale how emotionally intense they thought the scene was, with 1 = Not at all intense to 5 = 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 extent 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 pair¹.

Retrieval. Depending on their assigned condition, participants then either waited in the lab for a few minutes or left the lab for one day or one week, respectively. When they returned for the retrieval task, participants in the involuntary condition performed a simple visual attention task, while the 32 sounds from encoding and 32 filler sounds were played back to them twice, once to each ear. Sounds were replayed twice in order to ensure a sufficient number of memories in the involuntary condition, where memory activation is believed to be incidental. On each trial, a central fixation cross appeared for one second followed by a sound. During retrieval, each sound was panned 75% to the left or to the right and would therefore appear either slightly to the right or slightly to the left when heard through headphones. A bright yellow star appeared 1.5 seconds into the playback of the sound and remained visible until the participant gave a response. The star would be presented either to the right or to the left on the screen in an unpredictable pattern. Participants were asked to indicate in which side of the screen the star was located by pressing "1" for left and

¹ Ratings of relatedness were used to statistically control for the possible confound of semantic relatedness between sound cue and scene. As reported in Staugaard and Berntsen (2014), across the four experiments we found no consistent relationship between strength of narrative relatedness and frequency of memories (data available upon request).

“2” for right. The 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).

Participants in the involuntary condition were given misleading information to suggest that identifying the location of the star was the primary task. They were asked to indicate with a button-press, which side of the screen the star was in. In addition, they were told to indicate if they “suddenly and unexpectedly” experienced an image appearing in their minds and were told that the image could be anything. If they noticed an image suddenly appearing in their minds, they pressed “3”, which displayed a questionnaire on the screen. Participants were instructed to write 1-3 keywords describing the retrieved scene and then rated 1) the specificity of the retrieved scene, which was operationalized as whether the memory referred to a single, specific scene or was a mix of several scenes (from 1 = “Not specific at all” to 5 = Highly specific); 2) the emotional intensity of the retrieved scene (from 1 = “Not intense at all” to 5 = “Very intense”); 3) and their bodily reactions to the retrieved scene, such as increased heart rate, nervousness, or tension (from 1 = “Not at all to” 5 = “Very much”). Participants in the involuntary condition saw a total of 128 trials, (16 unique and 16 repeated sounds from encoding played twice, and 16 unique and 16 repeated filler sounds played twice).

Participants in the voluntary condition were asked to try to retrieve a scene from the encoding task in response to each sound. If they retrieved a scene, they were asked to describe it in a few keywords using the keyboard and then proceed with the task. If they were unable to retrieve a scene, they simply continued to the next trial.

Recognition. Participants were not told about the recognition task at any point before actually having to do it. All participants were shown each of the 32 scenes from encoding one after the other. Each scene was paired with a matched foil scene, and participants were asked to indicate which of the scenes was from encoding. They rated their confidence in their response on a 5-point Likert scale (1 = “Very uncertain” to 5 = “Very certain”).

Data preparation. In the original study by Staugaard and Berntsen (2014), scenes were either emotionally negative or neutral. In order to investigate the interaction between cue distinctiveness and retrieval condition over time, we excluded all the emotional memories from the present analysis for the following reasons. First, pictures of highly stressful scenes may not easily compare to everyday memories. Second, the memories of the emotional scenes were affected differently by retention time compared with the memories of neutral scenes in the original study (Staugaard & Berntsen, 2014), consistent with emotion having known effects on retention (e.g., McGaugh, 2004). We calculated the percentage of neutral memories in relation to the total number of cues for neutral scenes in each condition (voluntary = 16 cues, involuntary = 32 cues) and analyzed the memory frequency in the six groups (i.e., two retrieval conditions across three retention intervals).

Memories were coded in the following way: Correct memories were defined as keyword descriptions that corresponded to the scene that the sound cue was paired with during encoding. Errors were defined as memories recorded in response to a filler sound or a keyword description that did not match any of the scenes from encoding. Finally, if participants described a scene from encoding that had originally been paired with a different sound cue than the one they responded to, we coded this as an indiscriminately cued memory (for a definition of this term, see Staugaard & Berntsen, 2014). Errors were excluded from all analyses, but indiscriminately cued memories were included as valid responses. Since

repeated cues were very similar to one another in this design, it would not make sense to expect participants to be able to tell repeated cues apart (see also discussions in Berntsen et al., 2013; Staugaard & Berntsen, 2014). For this reason, we considered all repeated cues equivalent when coding accuracy (e.g., if a scene that had been paired with the sound of a dog barking, either of the eight sounds of dogs barking were considered valid cues for that scene during retrieval). However, this also meant that it was statistically easier for participants to retrieve a correct scene in response to a repeated cue (where they have more chances) compared with a unique cue (where they have only one chance). Therefore, main effects of cueing should be interpreted with caution. Importantly, the predicted interaction between cueing and retrieval condition can still be interpreted as the coding scheme was identical in both retrieval conditions.

Results

Proportion of correct memories as an effect of time. We analyzed the proportion of memories relative to the number of cues in each condition with a mixed ANOVA, where cueing (unique vs. repeated) was a within-subjects factor and retrieval condition (voluntary vs. involuntary) and retention interval (immediate vs. 24-hours vs. 1-week) were between-subject factors. We found a main effect of retention interval, showing reduced memory frequency over time, $F(2,122) = 22.37, p < .001, \eta_p^2 = .27$, immediate: $M = 41.7 (SD = 24.5)$, 24-hours: $M = 32.4 (SD = 20.6)$, 1-week: $M = 19.7 (SD = 15.4)$, see Figure 1. We found a main effect of cueing, which reflected a higher proportion of memories in response to unique cues compared with repeated cues, $F(1,122) = 22.04, p < .001, \eta_p^2 = .15$, unique cues: $M = 41.7 (SD = 31.0)$, repeated cues: $M = 26.0 (SD = 24.0)$. We also found a main effect of retrieval condition, with voluntary retrieval resulting in a greater proportion of memories compared with involuntary retrieval, $F(1,122) = 91.15, p < .001, \eta_p^2 = .43$, voluntary memories: $M = 48.6 (SD = 20.5)$, involuntary memories: $M = 19.1 (SD = 15.0)$, see Figure 1.

We then found an interaction between retrieval condition and retention interval, showing a greater reduction of memory frequency over time in the voluntary retrieval condition, $F(2,122) = 4.82, p = .010, \eta_p^2 = .07$, see Figure 1. An inspection of the figure suggested that that voluntary memories showed greater forgetting between the immediate and the 1-week retention interval. We therefore ran the analysis with only those two intervals, excluding the 24-hour interval. This analysis showed the predicted interaction between retrieval condition and retention interval, $F(1,92) = 9.51, p = .003, \eta_p^2 = .09$.

Finally, we found an interaction between cueing and retention interval, showing that unique cues were less effective as time passed, whereas repeated cues did not appear to lose their effectiveness over time, $F(2,122) = 11.70, p < .001, \eta_p^2 = .16$. No other effects were statistically significant ($ps > .132, \eta_p^2 < .019$).

Specificity of correct memories. Memory specificity was analyzed with a mixed ANOVA with cueing (unique vs. repeated sounds) as a within-subject factor and retrieval condition (voluntary vs. involuntary) and retention interval (immediate vs. 24 hours vs. one week) as between-subject factors. Consistent with previous work, there was a main effect of retention interval, as specificity ratings dropped between the shortest and the longest interval, $F(2,88) = 23.90, p < .001, \eta_p^2 = .35$, immediate: $M = 4.3 (SD = 0.6)$, 24 hours: $M = 4.1 (SD = 0.6)$, one week: $M = 3.2 (SD = 1.1)$. The main effect of cueing could be considered a statistical trend, with unique cues leading to more specific memories than repeated cues, $F(1,88) = 3.57, p = .062, \eta_p^2 = .04$, unique cues: $M = 4.1 (SD = 1.0)$, repeated cues: $M = 3.8 (SD = 1.1)$. In addition, we found an interaction between retrieval condition and retention interval, which was caused by a higher specificity for voluntary memories during the immediate interval, while there was no difference between groups during the 1-week interval,

$F(2,88) = 3.25, p = .044, \eta_p^2 = .07$, see Figure 2. No other effects were statistically significant ($ps > .177$, all $\eta_p^2 < .040$).

Response time of correct memories. Response time (in milliseconds) was analyzed with a mixed ANOVA with cueing (unique vs. repeated sounds) as a within-subject factor and retrieval condition (voluntary vs. involuntary) and retention interval (immediate vs. 24 hours vs. one week) as between-subject factors. We transformed the response times for the inferential statistics using a logarithmic transformation, but we report the untransformed descriptive data. Consistent with previous work, there was a main effect of retrieval condition with voluntary retrieval being slower than involuntary retrieval, $F(1,88) = 60.19, p < .001, \eta_p^2 = .41$, voluntary: $M = 7242 (SD = 2483)$, involuntary: $M = 3730 (SD = 1797)$. We also found an interaction between cueing and retention interval, showing that only memories in response to unique cues became slower over time, whereas memories in response to repeated cues were equally fast across the three retention intervals, $F(2,88) = 3.66, p = .030, \eta_p^2 = .08$. No other effects were statistically significant ($ps > .123$, all $\eta_p^2 < .047$).

Analysis of errors. We analyzed the proportion of indiscriminately cued memories in a factorial ANOVA with retrieval condition (voluntary vs. involuntary) and retention interval (immediate vs. 24 hours vs. one week) as between-subject factors. This analysis showed that the proportion of indiscriminately cued memories increased over time for all participants, main effect of retention interval: $F(2, 122) = 7.51, p = .001, \eta_p^2 = .110$. No other effects were significant ($ps > .170$, all $\eta_p^2 < .016$). We did not analyze memories in response to filler sounds, since there were no filler sounds in the voluntary condition.

Recognition. In order to investigate if differences between retrieval conditions might be due to differences in encoding, we included a surprise recognition task. We calculated the proportion of correct responses and mean ratings of confidence for each

participant and entered them into two separate factorial ANOVAs with retrieval condition (voluntary vs. involuntary) and retention interval (immediate vs. 24 hours vs. one week) as between-subject factors. The main effect of retention interval on accuracy could be considered a statistical trend, $F(2,122) = 2.61$, $p = .077$, $\eta_p^2 = .04$. No other effects of recognition accuracy approached statistical significance ($ps > .354$, all $\eta_p^2 < .018$). For recognition confidence, we found a main effect of retention interval showing a decline in confidence across the four intervals, $F(2,122) = 25.54$, $p < .001$, $\eta_p^2 = .30$, immediate: $M = 4.9$ ($SD = 0.2$), 24 hours: $M = 4.5$ ($SD = 0.4$), one week: $M = 4.2$ ($SD = 0.6$). No other effects of recognition confidence were statistically significant ($ps > .443$, all $\eta_p^2 < .006$).

Discussion

We analyzed data from a prior study (Staugaard & Berntsen, 2014) in order to investigate the relationship between cue distinctiveness and voluntary versus involuntary retrieval on memory accessibility over time. We found an interaction between retrieval condition and retention interval, which was the result of faster forgetting in the voluntary condition compared with the involuntary condition (Figure 1). Since involuntary memories at each delay approximated normal distributions, it is highly unlikely that a floor effect could influence the present results. Consequently, we argue that the interaction was caused by a reduced advantage of voluntary retrieval as time passed. We will return to this point in the General Discussion.

In addition to this novel finding, we replicated effects from earlier work, namely that voluntary retrieval was overall more effective than involuntary retrieval (Berntsen et al., 2013), unique cues were more effective than repeated cues (Berntsen et al., 2013; Hunt & Smith, 1996; Mäntylä & Nilsson, 1988), and involuntary retrieval was considerably faster than voluntary retrieval (Barzykowski & Staugaard, 2018; Berntsen et al., 2013; Schlagman & Kvavilashvili, 2008; Staugaard & Berntsen, 2014). We also found that retrieval time

increased with increased retention interval in response to unique cues, but not in response to repeated cues, which suggests that the unique associations may become less accessible than the repeated associations as time passes. We hypothesized that the advantage of voluntary retrieval would be particularly pronounced in response to repeated cues, but this prediction was not supported by the findings.

Finally, ratings of specificity were reduced after one week. This finding could reflect a tendency for memories to become schematized over time, which might also reduce their recollective quality. While the reduction in specificity happened in both retrieval conditions, it was more pronounced in the voluntary compared with the involuntary condition.

Some important limitations should be considered. It could be argued that some participants in the involuntary condition might use voluntary retrieval if they guessed the true purpose of the study. However, in Berntsen et al. (2013) we have evidence that introducing the parallel task in the voluntary condition interferes with voluntary retrieval by dramatically increasing response time. Even if participants wanted to use voluntary retrieval in the involuntary condition, this should impact their response times considerably. Given the lack of an interaction between retention interval and retrieval condition in relation to response time, it is highly unlikely that participants in either condition switched their assigned strategy.

In order to capture enough memories in the involuntary condition, participants assigned to involuntary retrieval were presented with twice as many cues as participants in the voluntary condition (see also Staugaard & Berntsen, 2014). Even though we normalized the memory frequencies by number of cues, a more ideal design would have similar numbers of cues in the two retrieval conditions. In addition, the original study (Staugaard & Berntsen, 2014) included emotional scenes in addition to the neutral scenes analyzed here. Although the emotional scenes were irrelevant to the present analyses, it is possible that exposure to these scenes and their retrieval also affected the retrieval of neutral memories to some extent.

Furthermore, Study 1 included only three retention intervals with a 1-week delay between Time 2 and Time 3, leaving performance at an intermediate interval unknown and thereby precluding more fine-grained analyses of forgetting. Finally, the lack of a parallel task in the voluntary condition meant that the retrieval tasks differed in their sensory demands between the two conditions, which might interfere with retrieval.

Study 2

The purpose of Study 2 was to remedy the shortcomings of Study 1. First, we changed the number of cues to be the exact same in the voluntary and involuntary condition, we included filler sounds in the voluntary condition, and we also administered the visual attention task in the voluntary condition. This meant that, as opposed to Study 1, the two retrieval conditions in Study 2 were completely identical save for the instructions given to participants. Second, we used scenes with a more positive valence than the ones used in Study 1 in order to increase their accessibility and reduce potential floor effects. Third, in order to plot retrieval success against well-established forgetting functions, we added an intermediate retention interval of 72 hours.

Methods

Participants. We recruited 134 students from Aarhus University and assigned them randomly to one of two retrieval conditions (voluntary vs. involuntary) and one of four retention intervals (immediate vs. 24 hours vs. 72 hours vs. one week). Sample size was based on prior studies using a similar design (Berntsen et al., 2013; Staugaard & Berntsen, 2014). Seven participants were excluded for not providing codable memories, leaving 127 participants for the analyses.

Materials. The encoding task consisted of 16 picture scenes from IAPS paired with 16 sounds. While the neutral scenes in Study 1 had been rather unremarkable (e.g., a clothes iron on an ironing board), we chose scenes for Study 2 that we judged to be more interesting

(e.g., a man drinking beer in a pub) and slightly positive according to the ratings that come with the IAPS set (mean valence = 5.6, range = 5.2 - 5.9; mean arousal = 3.9, range = 2.7 - 5.5). Each picture was rescaled at 1024x768 pixels and a mirrored version was created to be used in the recognition task. The sounds were a subset of those used in Study 1 and were either unique or repeated (see Appendix for the full list of cues and scenes). We also included 16 filler sounds divided into eight unique and eight repeated sounds. Two independent judges rated the semantic relatedness between sound cue and scene for each trial. Since the judges agreed (intraclass coefficient = .789), we ran a series of correlations between their mean rating of semantic relatedness for each of the 16 trials and the memory frequency for that trial. We ran these correlations separately for each participant in the immediate condition. Of the resulting 32 correlations, only one was statistically significant, 17 were negative, two were constants, and 13 were positive. We transformed the coefficients to z-scores, calculated an average, and transformed this value back to a correlation coefficient. Since the average correlation coefficient was .052, we conclude that semantic relatedness had no consistent effect on memory frequency (data available upon request).

The retrieval task was based on Study 1 and identical across all conditions. It consisted of 64 trials of a bright yellow star against a black background. A sound cue accompanied each presentation of the star.

The purpose of the recognition task was the same as in Study 1: We wanted to check if all participants had encoded the pictures. However, we made the surprise recognition task more difficult than in Study 1 to reduce the risk of a ceiling effect. The task consisted of each of the 16 pictures from the encoding task paired with its own mirror image. There were no obvious indicators in any of the pictures of the original orientation (e.g., identifiable letters).

Procedure. Participants were invited through a letter that described the study in general terms (e.g., “you will hear sounds and see pictures”), while avoiding any mention of

memory, forgetting, or memory accuracy. They were also informed that participation was completely voluntary and that they could withdraw their consent at any time. Participants arrived in the laboratory either by themselves or in small groups of up to four people. All members of a group were assigned to the same condition to conceal the purpose of the study.

The procedure was identical to the one described in Study 1, with the following exceptions: Participants were presented with only 16 trials during encoding. After presentation of a picture-sound pair, they were asked to describe the content of the picture using 1-3 keywords. After completing this task, participants saw the same 16 trials two times more in the same order as before in order to strengthen their associative learning between sound cues and scenes. Participants then either remained in the laboratory for a few minutes (i.e., the immediate retention interval), or left the laboratory for an interval of either 24 hours, 72 hours, or 1 week, depending on the condition they had been assigned to.

During the retrieval task, all participants were presented with the same 64 trials consisting of sounds played over the yellow star, as described in Study 1. The sounds were the 16 sounds from the encoding task (familiar sounds) replayed twice and an additional 16 sounds not previously heard (unfamiliar sounds) replayed twice for a total of 64 sounds. Unfamiliar sounds also consisted of eight unique and eight repeated sounds. Instructions for participants were identical to those in Study 1. Thus, the only two differences between the voluntary and the involuntary retrieval conditions were whether participants 1) were required to intentionally retrieve a scene in response to a cue or to wait for a scene to unexpectedly appear in their minds, and 2) whether they were required to perform the visual attention task by pressing the correct button or to simply watch the yellow star without responding.

During recognition, each picture from the encoding task was presented side by side with its mirror image on the screen at half the original size until response. Participants were instructed to indicate by button-press how the picture had been oriented during encoding and

rated their confidence in each of their responses. Upon completing the experiment, each participant received a gift card valued 100 DKK (approximately \$16).

Data preparation. All memory responses collected during the retrieval task were coded as either correct or incorrect memories. A correct memory was a response to a familiar sound cue where the keywords either 1) described the scene that the sound cue had originally been paired with during encoding, or 2) corresponded to the keywords given during the encoding task. Some participants might write idiosyncratic keywords like “MASH” or “Heath Ledger”, which did not readily identify a scene from encoding. However, if the participant had written keywords during the encoding task that included “MASH” or “Heath Ledger”, this indicated a correct memory. An incorrect memory was defined as 1) a description of a scene that had originally been paired with a different sound, 2) a response to a filler sound, or 3) a response that did not clearly describe a scene and did not correspond to any keywords from the encoding task (e.g., ‘holiday with Grandma’). While the first type of incorrect memory (i.e. indiscriminately cued memories) was included in the analysis for the same reasons as in Study 1, type 2 and 3 incorrect memories were excluded from the analyses.

Results

Frequency of correct memories as an effect of time. Since the number of cues was identical in all conditions, we analyzed the frequency rather than the percentage of correct memories as we did in Study 1. In order to meet assumptions of normality, we used a square root transformation on the total number of memories in each cue condition for each participant. However, we use the untransformed data when presenting the descriptive statistics. We entered the means into a mixed ANOVA with retrieval condition (voluntary vs. involuntary) and retention interval (immediate vs. 24 hours vs. 72 hours vs. one week) as between subject factors and cueing (unique vs. repeated cues) as a within subjects factor. The findings largely replicated Study 1. We found a main effect of retention interval,

demonstrating a drop in frequency of memories across the four intervals, $F(3,119) = 5.12$, $p = .002$, $\eta_p^2 = .11$, raw means for immediate: $M = 17.6$ ($SD = 9.5$), 24 hours: $M = 12.3$ ($SD = 9.0$), 72 hours: $M = 10.2$ ($SD = 6.6$), one week: $M = 8.4$ ($SD = 6.2$). We also found a main effect of cueing, with unique cues leading to a greater frequency of memories than repeated cues, $F(1,119) = 41.19$, $p < .001$, $\eta_p^2 = .26$, means for unique cues: $M = 8.1$ ($SD = 4.0$), repeated cues: $M = 6.0$ ($SD = 5.8$). Finally, we found a main effect of retrieval condition with participants in the voluntary condition showing a higher frequency of memories compared with the involuntary condition, $F(1,119) = 12.58$, $p = .001$, $\eta_p^2 = .10$, means for voluntary: $M = 14.5$ ($SD = 8.6$), involuntary: $M = 9.7$ ($SD = 8.0$). In contrast to Study 1, the interaction between retention interval and retrieval condition was not statistically significant, $F(3,119) = 0.86$, $p = .465$, $\eta_p^2 = .02$. However, inspecting the means (Figure 3), we felt it was justified to compare the immediate time interval with the 1-week interval, excluding the two intermediate retention intervals from the analysis. This analysis replicated the interaction observed in Study 1 (Figure 3): A greater number of voluntary memories were reported at the immediate interval, but there was no difference between the two conditions at the 1-week interval, $F(1,60) = 5.65$, $p = .021$, $\eta_p^2 = .09$. The interaction between cueing and retention interval found in Study 1 was not replicated in Study 2, $F(3,119) = 1.90$, $p = .134$, $\eta_p^2 = .05$. No other effects were statistically significant ($ps > .169$, all $\eta_p^2 < .017$).

In order to analyze retrieval success in greater detail in the two retrieval conditions, we plotted the total frequencies of correct memories in the two conditions as a function of delay and tried to fit standard retention functions to these distributions (Rubin & Wenzel, 1996). In both cases, a logarithmic function provided the best fit, accounting for more than 90% of the variance. The function for voluntary memories was $y = -25.1\ln(x) + 211.65$; $R^2 = 0.9227$. The function for involuntary memories was $y = -11.02\ln(x) + 148.14$; $R^2 = 0.9518$.

The difference in the parameters is consistent with a faster forgetting rate (i.e., steeper drop in frequency) in the voluntary conditions (see Figure 4).

Finally, since all sound cues were played twice during the retrieval task, we wanted to examine any potential effects this might have had on retrieval success in the two retrieval conditions. To this end, we tagged each memory as “first reporting” or “second reporting”. We then calculated the proportion of “second reporting” memories to the total number of memories for each participant and entered this value into a factorial ANOVA with retrieval condition and retention interval as between-subject factors. This analysis showed that participants in the voluntary condition had a larger proportion of “second reporting” memories (i.e., they more frequently reported the same scene memory twice) compared with participants in the involuntary condition, $M = 0.49$ ($SD = 0.20$) vs. $M = 0.40$ ($SD = 0.26$), $F(1,117) = 5.39$, $p = .022$, $\eta_p^2 = .04$. No other effects of this analysis were statistically significant ($ps > .672$, all $\eta_p^2 < .014$).

Specificity of correct memories. Self-rated specificity of the retrieved scenes was analyzed with a mixed ANOVA with cueing (unique vs. repeated sounds) as a within subject factor and retrieval condition (voluntary vs. involuntary) and retention interval (immediate vs. 24 hours vs. 72 hours vs. one week) as between subject factors. As in Study 1, we found a main effect of retention interval, demonstrating decreased specificity across the four intervals, $F(3,93) = 4.34$, $p = .007$, $\eta_p^2 = .12$, immediate: $M = 4.0$ ($SD = 0.8$), 24 hours: $M = 3.7$ ($SD = 0.9$), 72 hours: $M = 3.6$ ($SD = 0.9$), one week: $M = 3.2$ ($SD = 1.0$). There was also a main effect of cueing with unique cues leading to more specific memories than repeated cues, $F(1,93) = 24.78$, $p < .001$, $\eta_p^2 = .21$, unique cues: $M = 3.9$ ($SD = 0.9$), repeated cues: $M = 3.3$ ($SD = 1.2$). No other effects were statistically significant (all $ps > .203$, all $\eta_p^2 < .027$), see Figure 5.

Response time of correct memories. The response times were positively skewed, so we transformed them using a logarithmic transformation. We report the untransformed data in milliseconds when presenting descriptive statistics. Response times were analyzed with a mixed ANOVA with cueing (unique vs. repeated sounds) as a within subject factor and retrieval condition (voluntary vs. involuntary) and retention interval (immediate vs. 24 hours vs. 72 hours vs. one week) as between subject factors, replicating Study 1. There was a main effect of retrieval condition with voluntary retrieval being slower than involuntary retrieval, $F(1,93) = 36.73, p < .001, \eta_p^2 = .28$, voluntary: $M = 5844 (SD = 2913)$, involuntary: $M = 3329 (SD = 1321)$. The main effect of time was also statistically significant, $F(3,93) = 3.86, p = .017, \eta_p^2 = .10$. This analysis showed that all participants became slower over time regardless of condition, immediate: $M = 3878 (SD = 2056)$, 24 hours: $M = 4647 (SD = 2106)$, 72 hours: $M = 4295 (SD = 1902)$, 1-week: $M = 5489 (SD = 3662)$. The interaction between retention interval and cueing found in Study 1 was not replicated, $F(3,93) = 0.62, p = .980, \eta_p^2 < .01$.

As with the analysis of memory frequency, we wanted to examine if “first reporting” memories and “second reporting” memories differed in terms of their response times. We therefore calculated the response times separately for each category of memories. The data were analyzed with a repeated measures ANOVA with retention interval and retrieval condition as between-subject factors and repetition (“first reporting” vs. “second reporting”) as a repeated measure. This analysis showed that memories were retrieved faster during the second reporting than during the first, $M = 4146 (SD = 2100)$ vs. $M = 5226 (SD = 2974)$, $F(1,104) = 46.01, p < .001, \eta_p^2 = .31$. There was also an interaction with condition, which reflected that the difference in response time between the first and the second reporting of a memory was smaller in the involuntary condition relative to the voluntary condition, $F(1,104)$

= 9.09, $p = .003$, $\eta_p^2 = .08$. In addition, we of course replicated the main effects of retrieval condition and retention interval reported above. No other effects were statistically significant (all p s > .293, all $\eta_p^2 < .036$).

Analysis of errors. As in Study 1, we analyzed the proportion of indiscriminately cued memories in a factorial ANOVA with retrieval condition (voluntary vs. involuntary) and retention interval (immediate vs. 24 hours vs. 72 hours vs. one week) as between-subject factors. This analysis showed that the proportion of indiscriminately cued memories increased over time ($F(3, 119) = 2.88$, $p = .039$, $\eta_p^2 = .068$), but this increase was only apparent in the voluntary retrieval condition as evidenced by an interaction between retention interval and retrieval condition: $F(3, 119) = 2.93$, $p = .036$, $\eta_p^2 = .069$.

We analyzed memories in response to filler sounds using the same strategy as above and found a main effect of condition, showing a higher rate of memories in response to filler sounds in the involuntary condition, $F(3, 119) = 11.26$, $p < .001$, $\eta_p^2 = .086$. The main effect of retention interval could be considered a statistical trend, $F(3, 119) = 2.26$, $p = .084$, $\eta_p^2 = .054$.

Recognition. We calculated the proportion of correct responses for each participant as well as their confidence ratings and entered them into two separate factorial ANOVAs with retrieval condition (voluntary vs. involuntary) and retention interval (immediate vs. 24 hours vs. 72 hours vs. one week) as between subject factors. A main effect of time showed variations in recognition accuracy across the four intervals, $F(3,119) = 10.93$, $p < .001$, $\eta_p^2 = .22$, with a substantial decline from the immediate to the one week delay, immediate: $M = 91.0$ ($SD = 9.5$), 24 hours: $M = 79.6$ ($SD = 14.0$), 72 hours: $M = 83.9$ ($SD = 9.8$), one week: $M = 74.3$ ($SD = 13.4$). This result was also supported by a main effect of time showing a decline in confidence across the four intervals, $F(3,119) = 15.23$, $p < .001$, $\eta_p^2 = .28$, immediate: M

= 4.1 ($SD = 0.6$), 24 hours: $M = 3.9$ ($SD = 0.6$), 72 hours: $M = 3.7$ ($SD = 0.7$), one week: $M = 3.1$ ($SD = 0.6$). No other effects were statistically significant (all $ps > .213$, all $\eta_p^2 < .038$).

Discussion

Study 2 largely replicated the findings from Study 1 with some exceptions. The decrease in memory frequency over time was best accounted for by a logarithmic function for both types of memories (explaining more than 90% of the variance for both) with the parameters of the functions indicating a steeper slope (i.e., more rapid drop in reported memories) in the voluntary condition. This was consistent with results of analyses of variance. When examining the mean frequency of memories over time, we found that voluntary retrieval showed faster forgetting between the shortest and the longest retention interval, although this effect was only statistically significant, when we excluded the two intermediate retention intervals. The distributions of memories at the different retention intervals were approximately normal, which indicates that a floor effect is unlikely to account for the finding. We believe then that the interaction was caused by reduced efficiency of voluntary retrieval over time, and we will unfold this argument in the General Discussion.

We found that voluntary retrieval led to a higher frequency of memories compared with involuntary retrieval, but as in Study 1, and in contrast to our prediction, this advantage was not more pronounced in response to repeated cues. Participants in the voluntary condition were more likely to report the same memory twice compared with participants in the involuntary condition. This effect was consistent across the different retrieval intervals. It could reflect the strategic nature of voluntary retrieval, an argument we will return to in the General Discussion.

Consistent with Study 1, specificity ratings decreased over time, but, unlike Study 1, we did not find this to be more pronounced in the voluntary condition. We found a main

effect of cueing, with unique cues leading to more specific memories compared with repeated cues, consistent with a statistical trend in Study 1.

We also replicated the finding that involuntary memories are retrieved faster than voluntary memories. In Study 1, only uniquely cued memories became slower over time, whereas memories retrieved in response to repeated cues were equally fast after a week. However, in Study 2, we did not find this interaction. Instead, all memories became slower over time. Taken together, there is evidence of retrieval time increasing over longer retention intervals, but the effects are relatively small. Finally, the recognition task showed decreased accuracy and decreased confidence across the four retention intervals. Importantly, this finding was not influenced by retrieval condition, suggesting that the participants had encoded the scenes equally well, independent of their retrieval condition assignments.

General Discussion

We investigated the effects of the passage of time and cue distinctiveness on voluntary and involuntary memory for neutral scenes. Contrary to naturalistic studies, we found a faster rate of forgetting in the voluntary condition. The lack of an interaction between cue distinctiveness and retrieval condition makes it less likely that this increased rate of forgetting in voluntary memory can be explained by an increased vulnerability to interference. An alternative explanation is that the initial strategic advantage of voluntary retrieval disappears over time as retrieval success becomes more cue-dependent. Importantly, the difference in response time between retrieval conditions was maintained after a week, which indicates that participants in the voluntary condition were consistently using strategic retrieval processes, but also that these processes over time were no longer more effective than an associative and cue-dependent form of retrieval. One might argue that voluntary retrieval appears to be a waste of cognitive effort after longer delays, since it only works in the presence of relevant cues and since the same result can be achieved through involuntary

retrieval. Indeed, ratings of specificity were reduced over time equally much in both retrieval conditions, which suggests that the quality of the voluntary and involuntary memories was comparable. Involuntary retrieval was also consistently faster than voluntary retrieval, consistent with involuntary retrieval being less effortful.

We expected voluntary retrieval to lead to greater retrieval success in response to repeated cues than involuntary retrieval. However, this prediction was not supported in either study. This appears to challenge the theoretical notion that voluntary retrieval entails an increased capacity for interference resolution. Alternatively, the advantage of voluntary memory may reflect a more generalized ability to recall items and events from memory regardless of the associative strength of cues – or maybe even in the absence of cues altogether. The finding that participants in the voluntary condition in Study 2 were more likely to retrieve the same memory twice can be seen as tentative support for this hypothesis. In a study by Hall et al. (2014), the increased activity in dorsolateral prefrontal cortex during voluntary retrieval was apparent in response to both cues that were paired versus unpaired with scenes to be remembered. They concluded that the prefrontal activation reflects a sustained strategic search process rather than retrieval success, which may be seen as consistent with the present findings.

Finally, we found a reduction in recognition confidence in both studies, and a reduction in recognition accuracy in Study 2 (this effect was a statistical trend in Study 1). Importantly, these reductions were parallel in the two retrieval conditions, which indicates that the differences in voluntary and involuntary memory over time cannot be explained by differences in general memory ability or encoding success.

In short, the findings were highly consistent across the two studies. The two types of retrieval were similarly affected by cueing, and their specificity was similarly affected by retention time. In addition, both showed the standard forgetting curve for which a logarithmic

function provided a good fit accounting for more than 90% of the variance consistent with previous findings for involuntary and voluntary memories (Berntsen, 1998; Schlagman et al., 2009). The most important difference was a steeper decline in memory frequency in the voluntary condition, reducing the advantage of voluntary retrieval at longer retention intervals.

The finding that voluntary retrieval leads to a higher frequency of memories than involuntary retrieval during the shortest retention interval replicates earlier findings from our lab using a very similar paradigm (Berntsen et al., 2013; Staugaard & Berntsen, 2014). However, several studies investigating everyday autobiographical memory have found involuntary memories to be slightly more recent than voluntary memories (Barzykowski & Staugaard, 2018; Berntsen, 1998; Finnbogadóttir & Berntsen, 2011; Johannessen & Berntsen, 2010). One explanation for this apparent discrepancy could be that involuntary memories largely rely on environmental cues, while voluntary memories are less constrained by the immediate environment (Berntsen, 2009; Rasmussen, Johannessen, & Berntsen, 2014). Rasmussen et al. (2014) presented evidence suggesting that, when cueing conditions are similar, voluntary recall accesses at least as recent, and maybe even more recent, events as involuntary recall, perhaps because a general need probability constrains strategic search in favor of the most recent events (Anderson & Schooler, 1991). In turn, this could explain why the voluntary memories were more recent than involuntary memories in our study, given that the cueing conditions were identical.

In a broader perspective, the possibility that involuntary retrieval may be as effective as voluntary retrieval at longer delays for non-rehearsed memories is consistent with observations in literature that certain remote autobiographical memories may be activated involuntarily in response to a situational cue after having been out of reach through decades.

One frequently cited case is the French novelist Marcel Proust's (1928/1956) description of a vivid childhood memory elicited by the taste of a Madeleine cookie dipped in lime tea.

The unexpected activation of dormant memories in response to situational cues is also consistent with some observations in clinical psychology of 'recovered memories' of childhood trauma (e.g., Conway, 1997; Lindsay & Read, 1997, for reviews). Although the notion of recovered memories is contentious, and although the majority of such recovered memories appear to have been brought about through strategic retrieval attempts in the course of psychotherapy (e.g., Geraerts et al., 2007), there are some examples of recovered memories outside of therapeutic settings in response to situational cues (e.g., Bendiksen, 1997), which might be conceptualized as involuntary memories of forgotten events. However, the fact that the present studies used laboratory material without the personal significance and levels of complexity associated with real life events renders these possibilities highly tentative and speculative. Obviously, future research using longer retention intervals and more complex events would be needed to assess the relevance of the present findings in relation to such cases.

Conclusion

The findings were highly consistent across the two studies. Voluntary retrieval led to greater retrieval success than involuntary retrieval during the short retention intervals, and this advantage was independent of cue distinctiveness. At the same time, voluntary retrieval showed a faster rate of forgetting over one week. The present results indicate that the increased rate of forgetting cannot be explained by interference alone. Instead, we hypothesize that the associative strength between a cue and an item is the best predictor of memory performance at longer delays, while strategic retrieval has reduced effect. This entails that, in the presence of relevant and distinctive cues, involuntary retrieval may lead to

the same retrieval success as voluntary retrieval at longer delays despite being effortless and unintended.

Acknowledgements

The authors are grateful to Lisa Riisager, Heidi Kamp Lærke Madsen, and Nina Bylod Staugaard, who collected the data. The present research was supported by a grant from the Danish National Research Foundation (DNRF89; PI: Dorthe Berntsen), and a grant from the The Danish Council for Independent Research | Humanities (PI: Søren Risløv Staugaard).

References

- Anderson, J. R. & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, 2(6), 396-408. DOI:10.1111/j.1467-9280.1991.tb00174.x
- Barzykowski, K. & Staugaard, S. R. (2018). How intention and monitoring your thoughts influence characteristics of autobiographical memories. *British Journal of Psychology*, 109, 321-340. DOI:10.1111/bjop.12259
- Bendiksen, M. (1997). Circumstances and the Phenomenology of a Recovered Memory: A Corroborated Case Study. In D. S. Lindsay & J. D. Read (Eds.). *Recollections of Trauma: Scientific Evidence and Clinical Practice*, p. 566 (NATO ASI Series). New York: Plenum Press.
- Berntsen, D. (1996). Involuntary autobiographical memories. *Applied Cognitive Psychology*, 10(5), 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(2), 113-141. DOI:10.1080/741942071
- Berntsen, D. (2009). *Involuntary autobiographical memories. An introduction to the unbidden past*. Cambridge University Press.
- Berntsen, D. (2010). The unbidden past. Involuntary autobiographical memories as a basic mode of remembering. *Current Directions in Psychological Science*, 19(3), 138-142. DOI:10.1177/0963721410370301
- Berntsen, D. & Hall, N. M. (2004). The episodic nature of involuntary autobiographical memories. *Memory & Cognition*, 32(5), 789-803. DOI:10.3758/BF03195869

Berntsen, D., Staugaard, S. R., & Sørensen, L. M. (2013). Why am I remembering this now? Predicting the occurrence of involuntary (spontaneous) episodic memories. *Journal of Experimental Psychology: General*, *142*(2), 426-444. DOI:10.1037/a0029128.

Conway, M. A. (Eds) (1997). *Recovered memories and false memories*. Oxford: Oxford University Press.

Ebbinghaus, H. (1885/1962). *Memory: A contribution to experimental psychology*. New York: Dover.

Finnbogadóttir, H. & Berntsen, D. (2011). Involuntary and voluntary mental time travel in high and low worriers. *Memory*, *19*(6), 625-640.

DOI:10.1080/09658211.2011.595722

Geraerts, E., Schooler, J. W., Merckelbach, H., Jelicic, M., Hauer, B. J. A., & Ambadar, Z. (2007). The reality of recovered memories: Corroborating continuous and discontinuous memories of childhood sexual abuse. *Psychological Science*, *18*(7), 564-568. DOI:10.1111/j.1467-9280.2007.01940.x

Hall, S. A., Rubin, D. C., Miles, A., Davis, S. W., Wing, E. A., Cabeza, R., & Berntsen, D. (2014). The Neural Basis of Involuntary Episodic Memories. *Journal of Cognitive Neuroscience*, *26*, 2385-2399. DOI:10.1162/jocn_a_00633

Hunt, R. R. & Smith, R. E. (1996). Accessing the particular from the general: The power of distinctiveness in the context of organization. *Memory & Cognition*, *24*, 217-225. DOI:10.3758/BF03200882

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

Johannessen, K. B. & Berntsen, D. (2010). Current concerns in involuntary and voluntary autobiographical memories. *Consciousness and Cognition*, *19*(4), 847-860.

DOI:10.1016/j.concog.2010.01.009

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. Gainesville, FL: University of Florida.

Lindsay, D. S. & Read, J. D. (1997). *Recollections of Trauma: Scientific Evidence and Clinical Practice* (NATO ASI Series, 997).

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

Mace, J. H., Atkinson, E., Moeckel, C. H., & Torres, V. (2011). Accuracy and perspective in involuntary autobiographical memory. *Applied Cognitive Psychology, 25*(1), 20–28. DOI:10.1002/acp.1634

Mäntylä, T. & Nilsson, L. G. (1988). Cue distinctiveness and forgetting: Effectiveness of self-generated cues in delayed recall. *Journal of Experimental Psychology: Learning, Memory, Cognition, 14*, 502-509. DOI:10.1037/0278-7393.14.3.502

McBride, D. M. & Doshier, B. A. (1997). A comparison of forgetting in an implicit and explicit memory task. *Journal of Experimental Psychology: General, 126*(4), 371-392. DOI:10.1037//0096-3445.126.4.371

McBride, D. M., Doshier, B. A., & Gage, N. M. (2001). A comparison of forgetting for conscious and automatic memory processes in word fragment completion tasks. *Journal of Memory and Language, 45*(4), 585-615. DOI:10.1006/jmla.2001.2792

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

Proust (1928/1956). *Swanns way. In Remembrance of Things Past*. (vol. 1). Translated by C.K. Scott Moncrieff. New York: Random House.

Rasmussen, A. S., Johannessen, K. B., & Berntsen, D. (2014). Ways of sampling voluntary and involuntary autobiographical memories in daily life. *Consciousness and Cognition, 30*, 156–168. DOI:10.1016/j.concog.2014.09.008

Rubin, D. C. & Wenzel, A. E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review, 103*, 734-760. DOI:10.1037/0033-295X.103.4.734

Schlagman, S., Kliegel, M., Schulz, J., & Kvavilashvili, L. (2009). Differential effects of age on involuntary and voluntary autobiographical memory. *Psychology and Aging, 24*(2), 397-411. DOI:10.1037/a0015785

Schlagman, S., & Kvavilashvili, L. (2008). Involuntary autobiographical memories in and outside the laboratory: How different are they from voluntary autobiographical memories? *Memory & Cognition, 36*, 920-932. DOI:10.3758/MC.36.5.920

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

Staugaard, S. R., & Berntsen, D. (2014). Involuntary memories of emotional scenes: The effects of cue discriminability and emotion over time. *Journal of Experimental Psychology: General, 143*, 1939–1957. DOI:10.1037/a0037185

Tulving, E. & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning & Verbal Behavior, 5*(4), 381-391. DOI:10.1016/S0022-5371(66)80048-8

Wimber, M., Rutschmann, R. M., Greenlee, M. W., & Bäuml, K. H. (2009). Retrieval from episodic memory: Neural mechanisms of interference resolution. *Journal of Cognitive Neuroscience, 21*(3), 538-49. DOI:10.1162/jocn.2009.21043

Figure 1. Percentage of correct memories over time in the two retrieval conditions in Study 1. Error bars show one standard error.

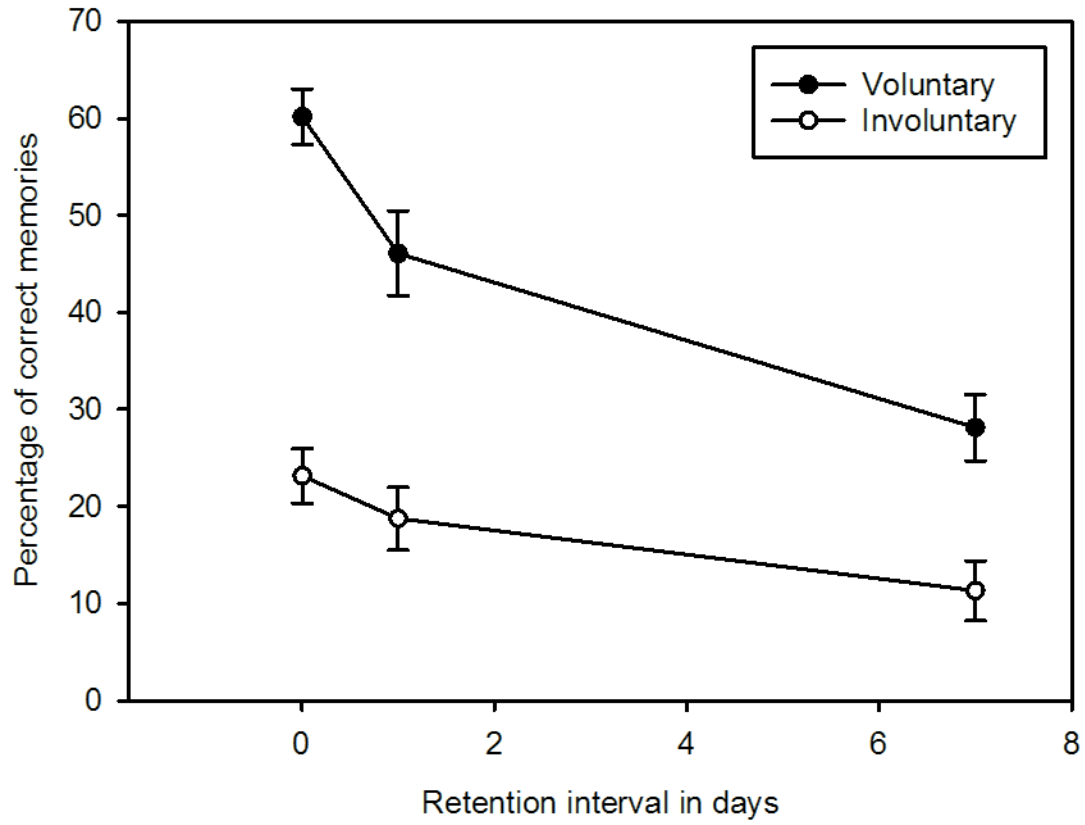


Figure 2. Specificity of memories in the two retrieval conditions over time in Study 1. Error bars show one standard error.

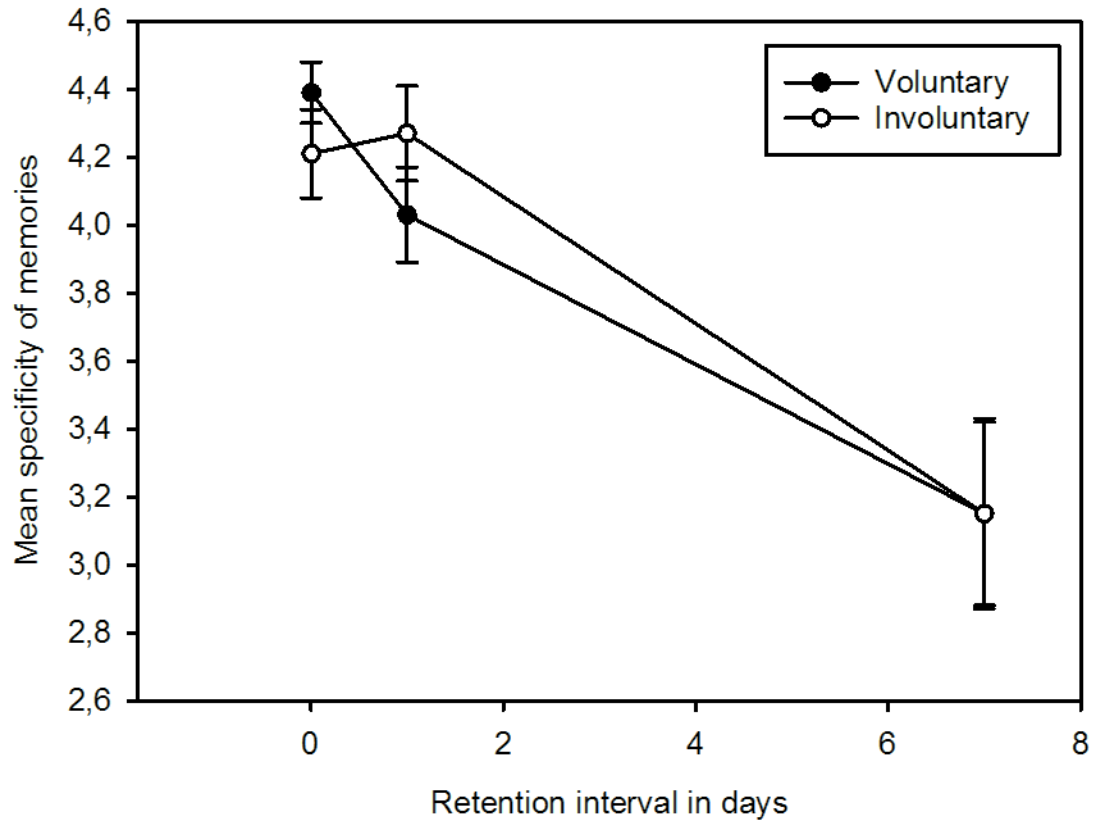


Figure 3. Mean frequency of correct memories over time in the two retrieval conditions in Study 2. Error bars show one standard error.

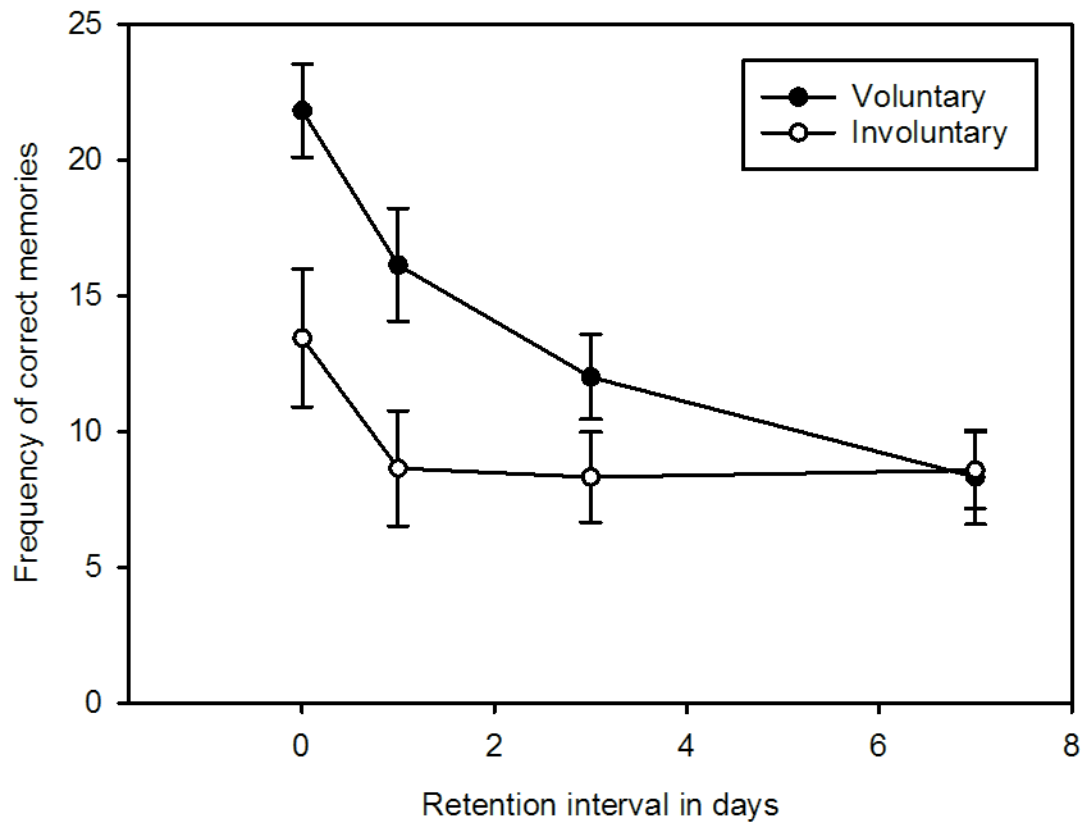


Figure 4. A logarithmic function fitting the retention of voluntary (top panel) and involuntary (bottom panel) memories across delays in Study 2.

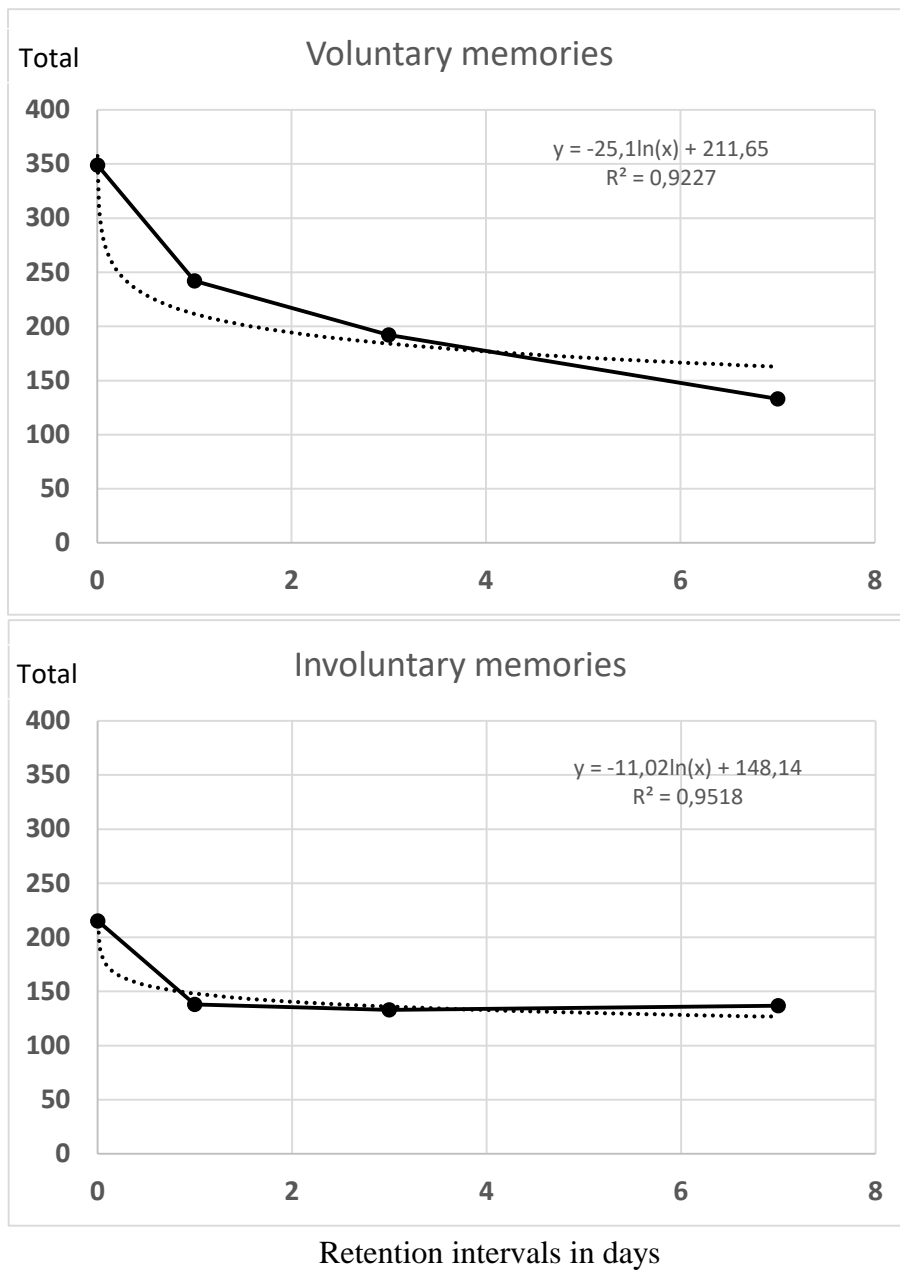
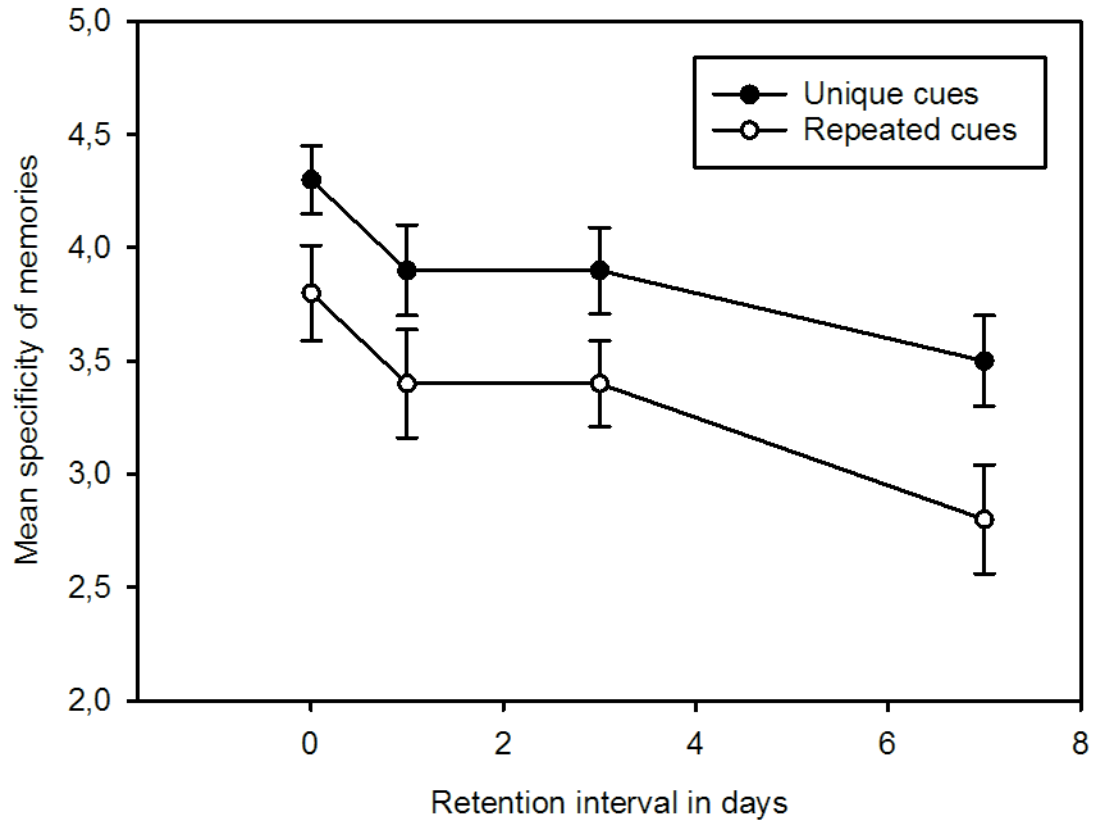


Figure 5. Specificity of memories in response to unique and repeated cues over time in Study 2. Error bars show one standard error.



Appendix. Cues and scenes in Study 2.

Cue type	Scene (IAPS code)	Sound
Unique	Sattelite dishes (5471)	Bell ringing
Repeated	Young woman on bed (2025)	Bird song 1
Unique	Woman playing golf (2506)	Chainsaw
Repeated	Chinese dragon (7600)	Bird song 2
Unique	Man holding newspaper (2102)	Music
Unique	Lizard on leaf (1121)	Phone
Repeated	Man with propeller (2575)	Bird song 3
Unique	Woman dancing (2606)	Car horn
Repeated	Rowing boats (5390)	Bird song 4
Repeated	Cowboy in snow (2635)	Bird song 5
Unique	Man kneeling on field (2191)	Trumpet
Repeated	Small cottage with flowers (5731)	Bird song 6
Repeated	Colorful squid (1947)	Bird song 7
Unique	Man holding beer (2600)	Dog barking
Repeated	Man with moustache (2020)	Bird song 8
Unique	Doctor holding journal (2394)	Camera