Notice: This is the author's version of a work that was accepted for publication in *Psychological Research*. A definitive version was subsequently published in *Psychological Research.* DOI: 10.1007/s00426-020-01304-5.

Through the Gateway of the Senses:

Investigating the Influence of Sensory Modality Specific Retrieval Cues on Involuntary Episodic Memory

Adam R. Congleton

Niels Peter Nielsen

Dorthe Berntsen

Center on Autobiographical Memory Research, Aarhus University, Denmark

Author Note

This research was supported by grants from the Danish National Research Foundation (DNRF89) and the Danish Council for Independent Research (DFF-6107-00050). Special thanks to Sinué Salgado and Daniel Munkholm Møller for helping manage data collection via Amazon Mechanical Turk.

Correspondence concerning this article should be addressed to: Adam R. Congleton, Center on Autobiographical Memory Research (CON AMORE), Department of Psychology and Behavioural Sciences, Aarhus University, Denmark. E-mail: congleton@psy.au.dk. Phone: +45 8716 6117.

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research*. A definitive version was subsequently published in *Psychological Research.* DOI: 10.1007/s00426-020-01304-5.

Abstract

Involuntary memories are memories of past events that come to mind with no preceding attempts of retrieval. They typically arise in response to situational cues, but little is known as to how such cues modulate involuntary memories. Here we examined how the sensory modality of the cues affects involuntary memory frequency and content. Participants watched first-person perspective films, and were later presented with visuospatial and/or auditory cues from the films. We then assessed their experience of involuntary memories for other moments from the films. Across Experiments 1 and 2, visuospatial cues resulted in a greater frequency of involuntary memories, and produced memories with a higher proportion of visual content. In Experiment 3, this effect was replicated using a more auditorily-engaging film and occurred whether participants focused on the film's auditory or visual components, but was more pronounced when there was a match between encoding fixation and the retrieval cue. These findings suggest visuospatial cues may outshine auditory cues in terms of involuntary memory elicitation and content.

Keywords: involuntary memory, episodic memory, retrieval cues, outshining effect, visual (modality) dominance effect

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research*. A definitive version was subsequently published in *Psychological Research.* DOI: 10.1007/s00426-020-01304-5.

Disclosure of Interest

The authors report no conflicts of interest.

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research*. A definitive version was subsequently published in *Psychological Research.* DOI: 10.1007/s00426-020-01304-5.

Through the Gateway of the Senses: Investigating the Influence of Sensory Modality Specific Retrieval Cues on Involuntary Episodic Memory

The majority of research on autobiographical/event memory has focused on examining how people consciously attempt to retrieve experiences from their past. However, sometimes people experience memories without the intention to do so. Such *involuntary memories* have been conceptualized as those that come to mind spontaneously without people consciously attempting to retrieve them (Berntsen, 1996). Although spontaneous retrieval is most often triggered by cues in the ongoing situation (Berntsen, 2009), involuntary memories are distinct from the standard concept of cued recall (e.g., Tulving & Pearlstone, 1966) by taking place with no conscious initiation of the retrieval process; the memories simply 'pop up' in response to accidental cues. The conceptual contrast to such involuntary episodic memories are memories of past events that are retrieved voluntarily – and, thus, come to mind as the result of a consciously-initiated and guided process (e.g., Berntsen, 1996). While involuntary episodic memories for many years were thought to be a special, primarily negative, class of memories associated with psychological disorders and trauma (e.g., Horowitz, 1975, 1986), a large body of more recent research has challenged this assumption. Specifically, involuntary memories have been found to occur as frequently as voluntary memories in our daily lives (Rasmussen & Berntsen, 2011), to be more frequently about positive than negative events, and to be affected by the same encoding mechanisms as voluntary memories (see Berntsen, 2009, 2012).

Evidence suggests that part of the reason involuntary memories emerge in the first place is the high degree of overlap that exists between an encoded event itself and some cue within the environment that reminds the person of that original event (Berntsen, 2009; also see Tulving's concept of synergistic ecphory; Tulving, 1984). However, one must also take into consideration the probable constraints upon the ability of retrieval cues to elicit involuntary memories. Specifically, if all cues we encountered in our daily lives evoked involuntary memories to the same degree, we would constantly be bombarded with memories of experiences from our past. One reason why such a situation does not occur is that the more items within a person's memory that are associated with any particular retrieval cue, the less effective that cue will be in terms of eliciting involuntary memories (i.e., the principle of cue overload; Watkins & Watkins, 1975; see Berntsen, Staugaard, & Sørensen, 2013 for an experimental demonstration of this effect with involuntary memories).

Many aspects of the nature of retrieval cues have been investigated in naturalistic diary studies, including whether cues that are effective at eliciting involuntary memories emerge primarily through external, internal, or through a combination of external and internal contexts (e.g., Berntsen, 1996; 2001; Berntsen & Hall, 2004; Mace, 2004; Schlagman, Kvavilashvili, & Schulz, 2007). A consistent finding has been the dominance of external cues (i.e., cues from the environment) over internal cues (i.e., thoughts and feelings; see Berntsen, 2009, for a review). Moreover, when the modality of reported sensory cues is compared, it has been found that visual/visuospatial cues are most frequently recorded, followed by auditory cues, whereas olfactory or tactile cues are rare (Berntsen, 2009). In addition, diary studies have also investigated the various categories of communalities that may exist between the retrieval cue participants report as triggering an involuntary memory and the content of the subsequent memory itself (such as specific objects, personal and impersonal life themes, activities, people, locations, or sensory experiences; Berntsen, 1996, 2001; Berntsen & Hall, 2004). Thus, it appears that the nature of retrieval cues and their connections with encoded events play an essential role in the elicitation of involuntary memories. Yet information about the nature of how different types of external retrieval cues affect the accessibility and sensory content of involuntary memories remains to be addressed. In addition, the findings from naturalistic studies have been based on self-reports. Experimental research is needed in order to draw firmer conclusions.

One question that remains unanswered and calls for experimental research is whether retrieval cues that are present simultaneously may compete with one another in the elicitation of involuntary memories, specifically when people encounter multiple cues within an environment that could all potentially re-activate their memory for an encoded event. Research on the *cue outshining effect* may offer some insight into this question. Specifically, it has been demonstrated that some environmental cues that have been encoded may fail to elicit memories because non-contextual information that has also been encoded, such as cues that involve inter-item associations, may be more effective at activating people's memory (e.g., McDaniel, Anderson, Einstein, & O'Halloran, 1989; see Smith, 2013, for a review). While cue outshining has been (primarily) studied with regard to the effects of contextual reinstatement on explicit and implicit tests of voluntary memory (e.g., Smith & Manzano, 2010; Smith & Handy, 2014; Smith, Handy, Hernandez, & Jacoby, 2018), its potential role in the elicitation of involuntary memory remains unexplored.

Within the context of involuntary memories, it is possible that when participants encounter an environment rich in sensory cues that could all potentially trigger a memory for a specific encoded event, some of the cues may be more effective at eliciting the involuntary memories, and thus may "outshine" other potential cues. In other words, if two different cues were

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research*. A definitive version was subsequently published in *Psychological Research.* DOI: 10.1007/s00426-020-01304-5.

available within the same ongoing situation, where both could involuntarily facilitate memory access to the same past event, one cue might be more strongly associated with the event, and, thus, be more likely or more effective at triggering the memory compared to the other cue. For example, one can imagine that encountering either the sight or the sound of fireworks on New Year's Eve might bring to mind different aspects of the previous year's celebration, but encountering the sight of the fireworks is ultimately more effective at triggering access to the memory than the sound. If each cue only triggers the retrieval of certain aspects of the past event, and if one cue is always more likely to trigger the memory, then the information that might be triggered by the other, less effective cue (the sound of the fireworks in this case) may remain unrecalled (e.g., the person might remember how they admired the fireworks from the balcony, but not remember that their dog was whining on the kitchen floor because of the noise). If the sounds were presented alone, with no visual input, then the resulting spontaneous memory construction might be different (e.g., the suffering of the dog might be remembered).

Here we assume that visuospatial cues may possess an advantage over other types of sensory cues in eliciting involuntary memories. This hypothesis derives from the results of naturalistic diary studies, wherein participants were more likely to report their involuntary memories as being triggered by visual cues compared to other types of sensory impressions (e.g., Berntsen, 2009). In addition, research has also demonstrated the importance of visual imagery in the construction of scenes we experience during an event memory (e.g., Rubin & Umanath, 2015), as well as its importance for our sense of reliving these events (e.g., Rubin & Berntsen, 2009), and our level of episodic specificity and phenomenological experience of such events (e.g., Aydin, 2018). As such, there may be a *visual dominance effect* observed when visual cues compete with cues from other sensory modalities (although see Mazzoni, Vannucci, & Batool, 2014 for evidence that visual information in the form of pictorial cues may be less effective at eliciting involuntary memories than verbal cues). Understanding how the sensory modality of retrieval cues affects involuntary memory frequency and content also has important implications for researchers investigating triggers of intrusive memories in patients with post-traumatic stress disorder (e.g., Ehlers, Hackmann, & Michael, 2004). Of particular relevance to the present study is the finding that visual sensations appear to be the most common form of intrusive symptom reported by such patients (e.g., Ehlers & Steil, 1995; Ehlers, Hackmann, Steil, Clohessy, Wenniger, & Winter, 2002).

In the process of investigating involuntary memories, and how retrieval cues influence their elicitation, a variety of different methods have been employed over the years. Such methods include the *diary method*, where participants are asked to keep a record of any time a memory of the past comes to mind involuntarily (see Mace, 2007; Berntsen, 2009 for reviews), and the *induction method*, in which attempts are made to elicit involuntary memories within laboratory-settings during a vigilance test using an associative-based retrieval procedure (e.g., Schlagman & Kvavilashvili, 2008). These methods measure the activation of the participants' personal memories and do not control encoding. In contrast, the *trauma film paradigm* (e.g., Holmes, Brewin, & Hennessy, 2004) and the *sound-scene paired-associate paradigm* (e.g., Berntsen, Staugaard, & Sørensen, 2013) both attempt to control the encoding and/or retrieval of involuntary memories using methods designed to act as analogues to how people experience such memories in their daily lives. However, all of these methods possess some degree of limitation, either in terms of problems controlling encoding and/or retrieval, or due to the restricted or non-naturalistic nature of the materials used in the context of the studies.

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research*. A definitive version was subsequently published in *Psychological Research.* DOI: 10.1007/s00426-020-01304-5.

The primary aim of the present study was to examine what influence the sensory modality of the retrieval cue has on the frequency and phenomenological content of involuntary episodic memories. In addition, we also wanted to investigate whether certain types of sensory cues would outshine others during retrieval; specifically, whether we would observe a dominance effect of visuospatial cues over auditory cues in the elicitation of involuntary memories. To address these questions, we employed an experimental method that simulates how involuntary memories are activated in people's daily lives. This method attempts to overcome some of the obstacles faced in previous research by enabling us to control the encoding and subsequent retrieval of an event through systematic manipulation of the event details present, while simultaneously using materials that more closely represent how events unfold in people's daily lives. Specifically, we used a paradigm that we had developed previously to study voluntary episodic memories; namely, the Simulated Event Paradigm (Congleton & Berntsen, 2019).

As part of the procedure, participants watch films we designed, specifically with the goal of manipulating memory for episodes. The films depict first-person perspective views of going through scenarios that represents the types of episodes participants might experience on a daily basis as if they were experiencing such events themselves (in this case, walking through a house and running through a forest). The participants are provided with overarching goals to keep in mind in order to ensure immersive experiences. In the present study, participants then complete a distracter task following the presentation of the films. After completing this distracter task, the participants are provided with retrieval cues that differ in terms of their sensory modality during a secondary task that varies depending upon the experiment and the condition (e.g., visuospatial vs. auditory cues). Finally, their experience of any involuntary memories of the film that came to their mind spontaneously during this secondary task is assessed, especially the frequency and phenomenological content of these memories.

Experiment 1

The aim of the first experiment was to investigate how the sensory modality of the retrieval cue influences the frequency of involuntary episodic memories (IEMs), as well as the content of the subsequent memories produced. To overcome some of the methodological obstacles faced by previous research, we used the Simulated Event Paradigm described in the introduction (Congleton & Berntsen, 2019). After watching a first-person perspective film and completing a distracter task, participants engaged in a secondary task that varied depending upon the type of sensory retrieval cues provided to them – visuospatial or auditory. By examining the influence of cue fragments taken from the film, we are able to determine (1) whether visual cues result in a greater frequency of memories when the number of cues is controlled, and (2) whether there is a general sensory-modality specificity effect present, wherein visuospatial cues result primarily in visual content and auditory cues in auditory content. It is possible that the content elicited by visuospatial cues may be primarily visual in nature because the visual modality of the encoded event "outshines" any auditory content that may also have been encoded during the event (resulting in a sensory-modality specificity effect), especially given the dominance of visual imagery in event memory (Rubin & Umanath, 2015). However, the content elicited by auditory cues may not be as restricted.

Given the results of naturalistic diary studies, in which participants report their involuntary memories as being triggered predominantly by visual cues, we predicted that visuospatial cues would be more effective at eliciting IEMs compared with auditory ones. Thus, in situations in which both auditory and visuospatial cues are available and able to prompt involuntary retrieval of aspects of the same event, visuospatial cues should outshine auditory ones. In addition, we also predicted a potential sensory-modality specificity effect in terms of the content of the IEMs produced in response to visuospatial cues, but not necessarily in response to auditory cues. In other words, we expected that participants would experience IEMs with primarily visual content in response to visuospatial cues, but had no firm predictions regarding auditory cues.

Methods

Participants and design. Participants were recruited from the Amazon Mechanical Turk online subject pool from English-speaking countries. They were informed that the purpose of this study was to gain a better understanding of how people comprehend details about an experience if they are asked to adopt a specific perspective during that experience. There were 200 participants (92 males; age ranging from 20 to 81) altogether who were randomly assigned to one of two conditions, with 100 participants per condition. This was a two condition, between-subjects design, with Cue Condition (Visual Cues Only vs. Auditory Cues Only) acting as the independent variable. This sample consisted only of participants who completed the entire study and who passed all manipulation checks (see below for details). The number of participants was comparable to previous research that captured IEM frequencies using an online paradigm (e.g., Berntsen, Rasmussen, Miles, Nielsen, & Ramsgaard, 2017).

Materials.

Online survey. The surveys were created using an online survey generation database, Qualtrics. The study was advertised on the Amazon Mechanical Turk online subject pool website (as described above), with a link to the Qualtrics survey posted on that website.

First-person perspective film. In order to assess participants' memory for an event they experienced, we used a novel film clip that depicted a first-person perspective view of walking through a computer-generated house. This film clip had been constructed for use in a previous study examining the influence of emotional variables on voluntary episodic memories (Congleton & Berntsen, 2019). The film was constructed using Garrys Mod programming software and Hammer Editor to build the virtual layout map. A recording was then made of a first-person perspective walkthrough of the computer-generated house. As the video clip had to be constructed from scratch, we were able to select exactly what event details would be present, and thus what details participants could possibly encounter, giving us control over the encoding of the event.

While watching the video, the participants were asked to imagine themselves in the role of the film's protagonist. They were given the following instructions: "You are about to see a video that shows a first person perspective view of walking through a house. ***While watching this video, imagine that you are a thief who is going through this house.*** You know that there is a safe with rare currency located somewhere inside the house and you are searching for it in the various rooms. You know that once you find the safe and take the currency, you will immediately leave the house." They then proceeded to watch a scenario in which the protagonist moved throughout the various rooms of the house. They started outside in the garden before entering the house through the French windows. They then explored the living room and bathroom, before entering the kitchen and hearing the telephone ringing. After proceeding into the utility room, where they subsequently heard an alarm beep, they moved upstairs to the second floor of the house and went into the office. Afterwards, they went into a child's bedroom before moving into the master bedroom, where they proceeded to open a safe and remove the currency. They then proceeded to the back door, where they encountered a man as they were exiting the house.

Distracter task – Shipley Institute of Living Scale 2 – Vocabulary Test. To reduce the likelihood of participants' rehearsing event details from the film clip, we had them complete the vocabulary portion of the Shipley Institute of Living Scale 2 as a distracter task (Shipley, 1940; Zachary, 1986). This task involved participants seeing a series of vocabulary words, and being asked to identify a synonym for the target word from among four alternatives (e.g., if the target word was "fortify," they should select the synonym "strengthen"). This task was chosen because it was clearly unrelated to the film and because it most likely engaged both visuospatial (seeing and reading words on the screen, locating synonyms) and auditory (subvocal processing of words) components of working memory (e.g., Baddeley, 2000).

Secondary tasks. To assess the frequency and phenomenological content of involuntary memory, we designed two secondary tasks that were to be performed by the participants, depending upon their condition, following the first-person perspective film and the distracter task. In the Visual Cues Only condition, participants were shown a series of seven visuospatial cues in the form of still images taken from the film they had earlier seen. They were told to simply watch the images as they were presented, with the survey auto-advancing after a set period of time. In the Auditory Cues Only condition, participants were presented with a series of seven auditory cues in the form of sounds taken from the film that corresponded exactly with the still images from the Visual Cues Only condition. For example, at one point in the film participants encounter a telephone ringing. In the Visual Cues Only condition, the participants saw the still image of the telephone taken from the video, while in the Auditory Cues Only condition they heard the sound of the telephone ringing, taken from the same moment in the video. The visual cues and the auditory cues were presented for the same length of time across both conditions.

The visuospatial and auditory cues provided to participants were not completely separated from the broader environmental context in which they were embedded due to the limitations of our computer software. The visuospatial cues consisted of a zoomed-in version with the sensory cue in the center of the screen. The auditory cues consisted of the primary sound acting as the cue as well as some minor noise from the broader auditory environment in the background (e.g., the footsteps the protagonist made at that moment). As such, both the visuospatial and auditory cues were (relatively) balanced in terms of their actual content because they were both still somewhat embedded within the broader sensory environment from which they were taken.

Procedure. After providing informed consent, as well as basic demographic information, the participants were asked to listen to a sound file and to identify the name of the animal sound they heard (as a method of ensuring that participants had their audio turned on during the experiment). The participants were then informed that they would be watching a film and were provided with the instructions described above. After watching the film, the participants were asked to answer a few questions about their emotional response to the film. They then completed the Shipley Vocabulary Task as a distracter. Following the completion of this task, the participants completed a secondary task where they were exposed to cues in terms of either still images or auditory clips. After completing this secondary task, participants were assessed concerning the frequency and phenomenological quality of any IEMs they experienced during this secondary task. Specifically, they were asked the following questions (all questions, excluding the Frequency question (#2), were assessed using a 7-point, Likert-type scale; see Table 1): (1) *Did seeing the pictures/listening to the audio clips at times spontaneously remind you of scenes that you had seen in the video (without you consciously trying to remember)?* (1 = Not at all, 7 = Almost constantly); (2) *Please indicate the approximate number of memories of*

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research*. A definitive version was subsequently published in *Psychological Research.* DOI: 10.1007/s00426-020-01304-5.

the video that spontaneously came to your mind (without you consciously trying to remember) while seeing the pictures/listening to the audio clips. (3) *To what degree were the spontaneous memories that came to your mind always of the same scenes (i.e., you had repeated memories of the same scene from the video while seeing the pictures/listening to the audio clips)?* (1 = They were never of the same scenes, 7 = They were always of the same scenes); (4) *When seeing the pictures/listening to the audio clips, the memories that came to my mind were predominantly visual (i.e., I saw them in my mind's eye).* (1 = Very rarely, 7 = Almost constantly); (5) *When seeing the pictures/listening to the audio clips, the memories that came to my mind were predominantly auditory (i.e., I could hear them in my mind).* (1 = Very rarely, 7 = Almost constantly); (6) *Did you react emotionally to the spontaneous memories that came to you while seeing the pictures/listening to the audio clips?* (1 = Not at all, 7 = Very strongly); (7) *Were the spontaneous memories that came to your mind while seeing the pictures/listening to the audio clips primarily positive or negative?* (1 = Primarily negative, 7 = Primarily positive).

This process of having the participants retrospectively report any instances of involuntary memories in response to prompts by an experimenter is a version of the probe-caught method of capturing spontaneous thought processes (e.g., see Maillet & Schacter, 2016 for a review). We used this particular method in the present study because it would have been difficult to allow participants to record such spontaneous thoughts online during the film sequences, which require active concentration. As previous research (e.g., Maillet & Schacter, 2016) has suggested that capturing spontaneous thoughts using the probe-caught method places less demands on thought monitoring processes in comparison to a self-caught method (where participants monitor their own thoughts and indicate the presence of involuntary memories as they come to mind), as well as resulting in a similar frequency of reported involuntary memories, this method was deemed

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research*. A definitive version was subsequently published in *Psychological Research.* DOI: 10.1007/s00426-020-01304-5.

most appropriate for the goals of this study. A version of this method was recently used by Berntsen, Rasmussen, Miles, Nielsen, and Ramsgaard (2017) in an online experiment, and also for that reason deemed feasible for the present study.

Manipulation Checks. As in Congleton and Berntsen (2019), where the Simulated Event Paradigm was also employed in a similar online-based format, a series of manipulation checks was included throughout the survey in an effort to ensure participants were keeping the overarching goals of the experiment in mind and following all instructions. If participants failed to answer any of the manipulation check questions correctly, or if they failed to correctly identify the animal sound they heard as part of the audio check, they were excluded from any further analysis. Thus, the results below are based only on data from participants who answered all manipulation and audio check questions correctly (i.e., 200 participants in total).

Results

In order to assess the influence of the sensory modality of retrieval cues on the frequency and phenomenological content of participants' involuntary memories, a series of one-way, between-subjects analyses of variance (ANOVAs) was conducted on the data. A Bonferroni correction was implemented whenever multiple comparisons were made.

General frequency of involuntary episodic memories. To assess the general frequency of IEMs that participants experienced while performing the secondary task, we examined their responses to the *Frequency* question on the phenomenological questionnaire (see Table 1). We then performed a square root transformation on the data to reduce the impact of extreme scores in the data set. The results of an analysis conducted on these transformed data revealed a significant effect of condition, where participants in the Visual Cues Only condition reported significantly more IEMs compared to those in the Auditory Cues Only condition (see Table 1).

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research*. A definitive version was subsequently published in *Psychological Research.* DOI: 10.1007/s00426-020-01304-5.

In addition to the results we obtained in terms of the *Frequency* of IEMs, we observed the same pattern of results when looking at the question regarding the *Occurrence* of IEMs (Table 1).

Influence of sensory modality of retrieval cue on phenomenological content of involuntary episodic memories. We next wanted to examine how the sensory modality of the retrieval cue affected the phenomenological content of IEMs. To address this question, we examined two questions from the phenomenological questionnaire in which we asked participants about the degree to which the retrieval cues they experienced during the secondary task resulted in IEMs that contained visual content or auditory content (i.e., the *Visual Content* and *Auditory Content* questions). As shown in Table 1, visual cues were associated with memories with more visual imagery and vice versa for auditory cues. In order to examine this relationship in greater depth, we conducted a 2 x 2 Mixed Model ANOVA, with Cue Type (Visual Cues vs. Auditory Cues) acting as the between-subjects factor and Sensory Content (Visual vs. Auditory) acting as the within-subjects factor (see Figure 1). The results revealed a main effect of Sensory Content, with participants reporting significantly more IEMs with Visual Content (M = 5.15) compared with Auditory Content (M = 3.71), $F(1, 198) = 79.37$, $p < .001$, $\eta^2_p = .29$. The results also revealed a significant interaction between Cue Type and Sensory Content, $F(1, 198) = 31.57$, $p < .001$, $\eta^2_p = .14$, where participants in the Auditory Cues Only condition reported experiencing IEMs containing an equal degree of visual and auditory content, whereas participants in the Visual Cues Only condition reported experiencing IEMs with significantly more visual content compared to auditory content (see Table 1 and Figure 1).

Discussion

The results of this experiment demonstrated that participants receiving visuospatial cues during the secondary task did indeed report significantly more IEMs compared to those who received auditory cues. Further, in terms of whether the sensory modality of the retrieval cue would affect the phenomenological content of the IEMs, the results indicated that memory content congruent with the sensory mode of the cues was more common for both visual and auditory cues. However, an interaction effect also showed that participants who received visuospatial cues were significantly more likely to report experiencing IEMs that contained a greater amount of visual than auditory content, whereas participants who received auditory cues reported experiencing IEMs that contained visual and auditory content to the same degree. Thus, we observed an effect of sensory-modality specificity only on the content of the IEMs elicited by the visuospatial cues. It appears as if visual information (in this film) not only has a dominating influence over auditory cues in eliciting IEMs in general, but it also to some extent outshined the retrieval of auditory content that was present during the event. In that sense, auditory cues may actually enable participants to retrieve a wider breadth of (sensory) details from an encoded event in comparison to visuospatial ones (a point to which we will return in the General Discussion).

Experiment 2

The results of Experiment 1 indicated that visuospatial cues were more powerful than auditory cues in relation to activating involuntary memories. Both visual and auditory memory content was most pronounced in response to sensory congruent cues; for example, more auditory contents were recalled in response to auditory cues than to visual cues. However, this advantage was not equally pronounced for the two types of cues. Memories retrieved in response to visuospatial cues had more visual than auditory imagery content, whereas no such difference was observed for auditory cues. Thus, it appeared that visuospatial cues were exerting a dominating influence over auditory cues in terms of affecting both the frequency and content of IEMs. However, in Experiment 1, we only examined the influence of visuospatial and auditory retrieval cues in isolation from one another. In Experiment 2, we wanted to both replicate and extend these findings by including an additional condition in which the visuospatial and auditory cues were presented in conjunction with one other (but, as in the previous experiment, isolated as much as possible from the overall surrounding context in which they originally occurred). The inclusion of this condition enabled us to determine whether the dominating influence of the visual modality we observed in Experiment 1 with regard to the content produced by visuospatial cues as compared to auditory cues would continue to be present when the cues were presented together, further supporting the idea of a general visual dominance effect.

Methods

Participants and design. As in the previous experiment, participants were recruited from the Amazon Mechanical Turk online subject pool from English-speaking countries. There were 320 participants (140 males; age ranging from 7 to 73) altogether who were randomly assigned to one of three conditions, with 99 participants in the Visual Cues Only condition, 100 participants in the Auditory Cues Only condition, and 121 participants in the Combined Cues condition. This was a three condition, between-subjects design, with Cue Condition (Visual Cues Only vs. Auditory Cues Only vs. Combined Cues) acting as the independent variable.

Materials.

Online survey. See Experiment 1 for details.

First-person perspective film. See Experiment 1 for details.

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research*. A definitive version was subsequently published in *Psychological Research.* DOI: 10.1007/s00426-020-01304-5.

Distracter task – Shipley Institute of Living Scale 2 – Vocabulary Test. See Experiment 1 for details.

Secondary tasks. The secondary tasks were exactly the same as in Experiment 1 with regard to the Visual Cues Only and Auditory Cues Only conditions, along with the addition of a third task known as the Combined Cues condition. In this third condition, participants were shown a series of seven cues that consisted of a combination of still images and sounds taken from the film they had earlier seen. These visual and auditory cues corresponded to the same moment in the film (i.e., they happened simultaneously). For example, at one point in the film, participants encountered a telephone ringing. In the Visual Cues Only condition, the participants saw the still image of the telephone taken from the film, while in the Auditory Cues Only condition they heard the sound of the telephone ringing, taken from the same moment in the film. In the Combined Cues condition, the participants both saw the still image and heard the sound of the telephone ringing simultaneously. All cues were presented for the same length of time across the three conditions.

Procedure. The procedure was identical to Experiment 1, except that during the secondary task participants were exposed to the Visual Cues Only, the Auditory Cues Only, or the Combined Cues conditions of retrieval cues. All other aspects of the procedure that occurred prior to and following the secondary task were the same as in Experiment 1.

Manipulation Checks. See Experiment 1.

Results

In order to assess the influence of the sensory modality of retrieval cues on the frequency and phenomenological content of participants' involuntary memories, a series of one-way, between-subjects analyses of variance (ANOVAs) was conducted on the data. A Bonferroni correction was implemented whenever multiple comparisons were made.

General frequency of involuntary episodic memories. To assess the general frequency of IEMs that participants experienced while performing the secondary task, we examined their responses to the *Frequency* question on the phenomenological questionnaire (see Table 2). As in Experiment 1, we performed a square root transformation on the data to reduce the impact of extreme scores in the data set. The results of an analysis conducted on these transformed data revealed a significant effect of condition (see Table 2), where participants in the Visual Cues Only condition reported more IEMs compared to those in the Auditory Cues Only, but not in the Combined Cues conditions. Follow-up pairwise comparisons revealed a significant difference between the Visual Cues Only and the Auditory Cues Only conditions ($p < .001$), and between the Auditory Cues Only and Combined Cues conditions ($p < .001$). As in the previous two experiments, in addition to the results we obtained in terms of the *Frequency* of IEMs, we also observed the same pattern of results when looking at the question regarding the *Occurrence* of IEMs (see Table 2).

Influence of sensory modality of retrieval cue on phenomenological content of involuntary episodic memories. As in the previous experiment, we next wanted to examine how the sensory modality of the retrieval cue affected the phenomenological content of IEMs. To address this question, we examined two questions from the phenomenological questionnaire, in which we asked participants about the degree to which the retrieval cues they experienced during the secondary task resulted in IEMs that contained visual content or auditory content (i.e., the *Visual Content* and *Auditory Content* questions), in order to further examine the relationship between the sensory modality of the retrieval cues and the content of the IEMs. We then conducted a 3 x 2 Mixed Model ANOVA, with Cue Type (Visual Cues vs. Auditory Cues vs. Combined Cues) acting as the between-subjects factor and Sensory Content (Visual vs. Auditory) acting as the within-subjects factor (see Figure 2). The results revealed a main effect of Sensory Content, with participants reporting significantly more IEMs with Visual Content (M = 5.22) compared with Auditory Content (M = 3.65), $F(1, 317) = 196.30$, $p < .001$, $\eta^2_p = .38$. The results also revealed a significant interaction between Cue Type and Sensory Content, $F(1, 317) = 24.36$, $p < .001$, $\eta^2_p = .13$.

To unpack this interaction term, we first calculated the difference in the reported sensory content of participants' involuntary memories by subtracting the auditory content from the visual content (i.e. *Visual Content – Auditory Content*). According to this calculation, positive scores would indicate that participants reported a greater amount of visual content compared to auditory content (and vice-versa for negative scores), while scores closer to zero would reflect a relatively equal amount of visual and auditory content. We then conducted a one-way between-subjects ANOVA on the data to determine if there were group differences regarding the amount of visual versus auditory content reported, with Cue Type (Visual Cues vs. Auditory Cues vs. Combined Cues) acting as the factor. The results revealed two findings. First, the means of all three groups were positive, indicating an overall greater amount of reported visual content compared to auditory content, although the effect was less pronounced in the Auditory Cues condition (see Figure 2). The results also revealed a main effect of Cue Type, with participants in the Visual Cues condition (M = 2.54) reporting significantly more visual content compared to participants in both the Auditory Cues (M = .56) and Combined Cues conditions (M = 1.62), $F(1, 317) = 24.36$, $p < .001$, $\eta^2_p = .13$. Follow-up pairwise comparisons revealed significant differences

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research*. A definitive version was subsequently published in *Psychological Research.* DOI: 10.1007/s00426-020-01304-5.

between the Visual Cues and Auditory Cues conditions ($p < .001$), the Visual Cues and Combined Cues conditions ($p = .002$), and the Auditory Cues and Combined Cues conditions ($p < .001$).

Discussion

In this experiment, we followed up on the results of Experiment 1 by examining the influence of the sensory modality of a retrieval cue on the frequency and content of IEMs reported by participants. In order to assess the possibility of visuospatial cues outshining auditory ones with regard to IEM content, we included a Combined Cues condition in addition to the Visual Cues Only and Auditory Cues Only conditions. The results replicated and extended the findings of Experiment 1, demonstrating once again a dominating effect of visuospatial cues over auditory ones in terms of the frequency and content of IEMs experienced by the participants. We also found that visuospatial cues were demonstrating a sensory-modality specificity effect in terms of IEM content, wherein they were more likely to produce IEMs with primarily visual content than with auditory content, whereas auditory cues were likely to produce IEMs with relatively equal amounts of visual and auditory content. Importantly, we further demonstrated that visuospatial cues were outshining the auditory ones in terms of both frequency and imagery content of IEMs, as they were equally likely to produce IEMs when presented either in isolation from or in conjunction with the auditory cues. We also observed that when visuospatial and auditory cues were presented simultaneously, the content of the resultant memories resembled the content evoked when visuospatial cues were presented in isolation. Auditory cues, on the other hand, yielded the most auditory content when presented alone. As such, we can say that the visual modality appears to be outshining the auditory modality of encoded events in two separate ways during retrieval: (1) at the level of the general frequency of IEMs being reported, and (2) at the level of phenomenological content.

Experiment 3

The previous two experiments demonstrated a dominating effect of visuospatial cues on the frequency (as well as the content) of IEMs. However, some questions remain unanswered. First, would the visual dominance effect continue to be present when a different movie is used, particularly if the movie is more "auditorily-engaging" than the house video? Second, would this visual dominance effect remain if participants were given a goal that encouraged them to focus more attention on the auditory components of the film as compared to the visual components during encoding? We examined these questions in Experiment 3, where we had participants watch a new first-person perspective film that was rated as being more auditorily-engaging than the house video (see Methods for details). Importantly, we asked participants to watch the video while keeping an overarching goal in mind that encouraged them to focus their attention on either the visual or auditory components of the film. After watching the film and completing the distracter task, participants then engaged in a secondary task that consisted of cue fragments taken from this new film. Finally, their experience of the frequency and phenomenological content of IEMs was assessed as in the previous experiments. Given the consistently dominating influence of visual information observed so far, we predicted that participants who received visuospatial cues would once again report a higher frequency of IEMs compared to those who received auditory cues. However, that effect may be tempered if participants are asked to focus their attention upon the auditory components of the film during encoding.

Methods

Participants and design. As in the previous experiments, participants were recruited from the Amazon Mechanical Turk online subject pool from English-speaking countries. This was a 2 (Encoding Fixation: Visual Fixation vs. Auditory Fixation) x 3 (Cue Type: Visual Cues Only vs. Auditory Cues Only vs. Combined Cues) between-subjects design. There were 602 participants (268 males; age ranging from 18 to 82 years) altogether who were randomly assigned to one of six conditions. There were 100 participants in the Visual Fixation-Visual Cues condition, 102 in the Auditory Fixation-Visual Cues condition, 101 in the Visual Fixation-Auditory Cues condition, 100 in the Auditory Fixation-Auditory Cues condition, 99 in the Visual Fixation-Combined Cues condition, and 100 in the Auditory Fixation-Combined Cues condition.

Materials.

Online survey. See Experiment 1 for details.

First-person perspective film. In order to determine if the visual dominance effect observed in the previous experiments would continue to be present when participants viewed a more auditorily-engaging film, we used a film clip that depicted a first-person perspective view of running through a forest. As with the house video used in the previous experiments, this film clip had been constructed for use in prior studies using the Simulated Event Paradigm (e.g., Congleton & Berntsen, 2019). We used the included editor software 'Eden' from the commercial video game 'ARMA III' to construct and script the virtual environment and its seven distinctive events. The playthrough of the constructed map was then recorded using an in-game recorder called 'Nvidia ShadowPlay'. For final editing of the raw recordings, we used a video editing software called 'Movie Studio'. As with the house video, a recording was then made of a first-person perspective walkthrough of the computer-generated forest.

While watching the video, the participants were asked to imagine themselves in the role of the film's protagonist (i.e., the runner). All participants were then given the following instructions: 'You are about to see a video that shows a first-person perspective view of running through a forest. While watching this video, imagine that you are the person who is running through this forest. You have been on vacation and are taking one last run along the path near your hotel.' For those participants in the Visual Fixation conditions, they were also told: 'You are very interested in birds, and you are looking for a rare bird that is hard to see that is supposedly found along this path.' For those participants in the Auditory Fixation conditions, they were also told: 'You are very interested in birds, and you are listening for a rare bird that is hard to hear that is supposedly found along this path.' They then proceeded to watch a scenario in which the protagonist ran along a path through the forest. Along the way, the protagonist passed a bicyclist, a group of cows, a man sitting at a log camp, a wind turbine, a tractor, another runner, and a group of seagulls flying overhead. All of these scenarios also had sounds that accompanied them (i.e., the sound of the bicycle, the cows mooing, etc.).

In order to determine if this video was more auditorily-engaging than the house video, we conducted an online ratings experiment prior to conducting Experiment 3. A total of 75 participants were drawn from the Amazon Mechanical Turk online subject pool from English-speaking countries. This rating experiment was a three-condition, between-subjects design in which participants were asked to watch the House Video, the Forest Video, or a third video that was not relevant for this particular study. They were asked to assess the video they saw in terms of several sensory qualities, including its Visual Interest, Visual Attention, Visual Repetitiveness, Auditory Interest, Auditory Attention, and Auditory Repetitiveness. The results indicated that the Forest Video was rated as significantly more Auditorily Interesting compared to the House Video ($p < .001$) and required more Auditory Attention ($p = .01$), while there was no difference between the two videos in terms of their Visual Interest ($p = .12$) or in terms of the amount of Visual Attention they required ($p = .66$). It was not considered feasible to design a video that was less visually-engaging and more auditorily-engaging, as such a situation is not something that typically happens in participants' daily lives (except in instances in which there is very limited visual information available, such as when one attends a concert or listens to audio books).

Distracter task. See Experiment 1 for details.

Secondary tasks. As in Experiments 1 and 2, the secondary tasks consisted of a series of cue fragments taken from the film, except in this case they were taken from the forest video (rather than the house video). In the Visual Cues Only condition, participants were shown a series of seven visuospatial cues in the form of still images taken from the film they had earlier seen. They were told to simply watch the images as they were presented, with the survey auto-advancing after a set period of time. In the Auditory Cues Only condition, participants were presented with a series of seven auditory cues in the form of sounds taken from the film that corresponded exactly with the still images from the Visual Cues Only condition. In the Combined Cues condition, participants were shown a series of seven cues that consisted of a combination of still images and sounds taken from the film they had earlier seen. These visual and auditory cues corresponded to the same moment in the film. For example, at one point in the film, the participants run past a group of cows. In the Visual Cues Only condition, the participants saw the still image of the group of cows taken from the film, while in the Auditory Cues Only condition they heard the sound of the cows mooing, taken from the same moment in the film. In the Combined Cues condition, the participant both saw the still image and heard the sound of the cows simultaneously. All cues were presented for the same length of time across the three conditions.

Procedure. The procedure was identical to the previous experiments, except that prior to watching the film the participants were given instructions that encouraged them to focus their attention on either the visual or auditory information while watching the film (see above for details). All other aspects of the procedure were the same as in Experiments 1 and 2, but with retrieval cues taken from the new film. Also, we reduced the number of questions we asked participants in the phenomenological questionnaire, retaining only the *Occurrence*, *Frequency*, *Visual Content*, and *Auditory Content* questions (as they were the only ones pertinent to this study).

Manipulation Checks. See Experiment 1 for details.

Results

In order to assess the influence of the sensory modality of retrieval cues and the encoding fixation instructions on the frequency and phenomenological content of participants' involuntary memories, a series of 2 x 3 between-subjects analyses of variance (ANOVAs) was conducted on the data. A Bonferroni correction was implemented whenever multiple comparisons were made.

General frequency of involuntary episodic memories. To assess the general frequency of IEMs that participants experienced during the secondary task, we examined their responses to the *Frequency* question on the phenomenological questionnaire (see Table 3). As in the previous experiments, we first examined the data for the presence of extreme scores that might affect the results. We discovered four outliers across the approximately 602 data points that were identified as statistically significant. After removing those four outliers from the entire data set, we then performed a square root transformation on the data in a further effort to reduce the impact of extreme scores in the data set, and to make the frequency means comparable across the three studies. We then conducted a 2 (Encoding Fixation: Visual Fixation vs. Auditory Fixation) x 3 (Cue Condition: Visual Cues Only vs. Auditory Cues Only vs. Combined Cues) between-subjects ANOVA on the transformed data. The results revealed a marginally significant effect of Cue Type. Follow-up pairwise comparisons revealed that participants in the Visual Cues Only conditions reported significantly more IEMs compared to those in the Auditory Cues Only conditions ($p = .05$), while there were no differences between the Visual Cues Only or Auditory Cues Only conditions and the Combined Cues condition (ps > .39).

The results of the ANOVA also revealed a significant interaction between Encoding Fixation and Cue Type, while there was no main effect of Encoding Fixation alone ($F = .26$). When unpacking the interaction, we observed that participants in Visual Fixation-Visual Cues condition reported the highest frequency of IEMs, whereas participants in the Visual Fixation-Auditory Cues condition reported the lowest (see Table 3). In contrast, there were no differences in the reported frequencies of IEMs for participants in the Auditory Fixation-Visual Cues, Auditory Fixation-Auditory Cues, or Auditory Fixation-Combined Cues conditions. In other words, visual cues were stronger but more so in combination with a visual focus during encoding.

As in the previous experiments, we also observed a lower frequency of auditorily-cued memories when looking at the *Occurrence* question of the phenomenological questionnaire (see Table 3).

Influence of sensory modality of retrieval cue on phenomenological content of involuntary episodic memories. We next wanted to examine how the sensory modality of the retrieval cue and the encoding fixation instructions affected the phenomenological content of

IEMs. To address this question, we examined two questions from the phenomenological questionnaire in which we asked participants about the degree to which the IEMs they reported experiencing contained visual and auditory content. We conducted a 2 x 3 x 2 Mixed Model ANOVA, with Encoding Fixation (Visual Fixation vs. Auditory Fixation) and Cue Type (Visual Cues Only vs. Auditory Cues Only vs. Combined Cues) acting as the between-subjects variables, and Sensory Content (Visual vs. Auditory) acting as the within-subject factor (see Figure 3). The results revealed a main effect of Sensory Content, with participants reporting significantly more IEMs with Visual Content ($M = 5.59$) compared to Auditory Content ($M = 4.66$), $F(1, 592) = 112.74$, $p < .001$, $\eta^2_p = .16$. There were no main effects of Encoding Fixation or Cue Type alone (all Fs < 2.87) and no significant two-way or three-way interactions among Sensory Content, Encoding Fixation, and Cue Type (all Fs < 2.7).

Discussion

In this study, we wanted to determine whether the visual dominance effect we observed in the previous experiments would replicate when a more auditorily-engaging video was used and when participants were encouraged to focus their attention on either the visual or auditory aspects of the video (during encoding). In terms of their influence on the frequency of IEMs, we found that visuospatial cues resulted in the highest reported number of IEMs, replicating the previous experiments, albeit with somewhat reduced effects. Visuospatial cues were also once again found to be just as effective at eliciting IEMs when presented alone as when presented in conjunction with the auditory cues. However, in this experiment we found that auditory cues were also equally effective at eliciting IEMs when presented by themselves or in conjunction with visuospatial cues.

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research*. A definitive version was subsequently published in *Psychological Research.* DOI: 10.1007/s00426-020-01304-5.

At the same time, we observed that the sensory fixation with which participants encoded the video also had an influence on the numbers of IEMs experienced. Specifically, participants who were asked to focus on the visual components of the film reported the highest number of IEMs when subsequently presented with visuospatial cues, whereas participants who were asked to focus on the auditory components reported the highest number when presented with auditory cues. However, if participants were presented with combined cues then it did not matter what their attention was focused upon during encoding. Across all conditions, irrespective of cuing, visual content was most salient in memory.

General Discussion

Across three studies, we examined how the sensory modality of the retrieval cue (and sensory fixation during encoding) would influence the frequency and phenomenological content of participants' involuntary episodic memories of a first-person perspective film they experienced. We investigated this question using a paradigm we had recently developed that attempted to overcome some of the obstacles faced by previous research on everyday involuntary memories, one that controlled both encoding and retrieval, while simultaneously using materials that more closely represented how events unfold in participants' daily lives (Berntsen et al., 2017).

The results revealed two important phenomena related to visuospatial cues: their relative dominance in eliciting involuntary memories over auditory cues, and their tendency to result in involuntary memories with primarily visual content. It appears that visuospatial cues are very effective elicitors of involuntary memories, as they were equally likely to produce IEMs when presented in isolation from auditory cues as when they were presented in conjunction with such cues, whereas auditory cues were most likely to elicit memories when presented together with visual cues than when presented alone. At the same time, the presentation of visuospatial cues primarily triggered visual content, resulting in a sensory-modality specificity effect that entailed a relative sparsity of content from the encoded event compared to auditory cues. This effect was also found when visuospatial cues were presented in conjunction with auditory cues. Here, the same pattern of results emerged as if the auditory cues were not present at all. These findings demonstrating the dominating effect of visual cues over auditory cues is consistent with more observational findings reported in diary studies of everyday involuntary memories, where it has been observed that visual cues are the most frequently recorded sensory modality (Berntsen, 2009). This pattern of results held true even when using a video that was more auditorily-engaging (i.e., the forest video in Experiment 3), although the effects were not as clear. However, the results of the third experiment demonstrated that visual cues resulted in fewer involuntary memories when participants were asked to adopt an auditory fixation during encoding. This finding is in line with classic research on transfer appropriate processing (e.g., Morris, Bransford, & Franks, 1977), and reveals that the effectiveness of visual cues at triggering involuntary memories is diminished when there is a mismatch between the fixation/orienting task used during encoding and the cue type encountered during retrieval.

Auditory cues seem to result in fewer involuntary memories in general, and they are most effective at producing such memories when presented together with visuospatial cues (except in cases where a more auditorily-engaging video/event is used, as in Experiment 3). However, while they may result in fewer involuntary memories in general, they are likely to result in IEMs that contain an equal amount of visual and auditory content (at least in situations with less distinct/engaging auditory information). Thus, such cues seem to result in a greater diversity of sensory content for a previously-experienced event. Taken together, these results seem to suggest that visuospatial cues may outshine the "elicitation potential" of cues from other sensory modalities, which may consequently constrain what aspects of previous events we may be able to re-experience involuntarily. Of course, more research is needed to replicate and extend the present findings by investigating other sensory modalities and by examining these phenomena using different materials. One immediate implication that emerges from these results is the fact that future research investigating involuntary memories should take into consideration the modality of the cues used to elicit such memories and the goal/instructions with which participants encoded the events, as both the frequency and content of the resultant memories may vary depending upon such factors (Mazzoni et al., 2014).

Although speculatively, our results may also have implications for clinicians treating patients with post-traumatic stress disorder. In the clinical literature, researchers have devoted attention to understanding the mechanisms of elicitation involved in the intrusive re-experiencing of traumatic events. Research has especially focused on the role of sensory cuing and imagery in relation to the trauma, although the majority of this research has tended to be observational in nature. In the process, it has been discovered that visual sensations seem to be the most frequently reported intrusive symptom by PTSD patients (e.g., Ehlers & Steil, 1995; Ehlers, Hackmann, Steil, Clohessy, Wenninger, & Winter, 2002). Researchers have found that these intrusive symptoms are often triggered by external cues that have a high degree of perceptual similarity to details surrounding the original traumatic event.

Based on the present findings, an advantage for visual cues and a dominance of visual content would not be specific to stressful or traumatic events. The results of the present study provide experimental evidence demonstrating the dominating influence that visual cues (and visual fixation) have on the frequency of reported involuntary memories, as well as the subsequent content of those memories. Here we observed that visuospatial cues may outshine the ability of auditory cues to elicit additional content of an encoded event when both cues are presented simultaneously. This visual dominance effect may help explain why PTSD patients report a greater degree of visual content in their intrusive memories. It may be the case that visual cues prompting intrusive memories may be suppressing the elicitation of other details from the event, including details that might aid the patients' recovery. Specifically, research has shown that enabling patients to elaborate upon their traumatic experience and place the intrusive triggers within the overall context of the event, as well as helping to properly identify all the triggers in general, is essential to helping ameliorate PTSD symptoms (e.g., Ehlers, Hackmann, & Michael, 2004).

In spite of the strengths of this study, there are a number of limitations. First, we used a probe-caught method to measure the frequency of involuntary memories (e.g., see Maillet & Schacter, 2016 for a review). While this method may be more suitable than the self-caught method, especially during a short experimental trial like the one used in the present experiments, it requires retrospection, which may introduce noise. However, any noise introduced by this recording technique most likely would have had similar effects across the different modality conditions and thus not change the basic findings. Still, future research should aim at replicating the present findings using other methodologies.

Second, our use of Amazon Mechanical Turk and online surveys in general might be seen as a potential limitation. However, research has demonstrated that the use of experimental methods conducted online under most conditions are as reliable and effective as those conducted within a laboratory environment (e.g., Grysman, 2014; Grysman, Prabhakar, Anglin, & Hudson, 2015; Berntsen, Rubin, & Salgado, 2015; see Mason & Suri, 2011 and Buhrmester, Talaifar, &

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research*. A definitive version was subsequently published in *Psychological Research.* DOI: 10.1007/s00426-020-01304-5.

Gosling, 2018 for reviews). In addition, online studies make it easier to involve considerably larger participant samples than what is typical in more conventional laboratory studies, which helps protect against the observation of random effects. While, of course, we do not have direct observation over the participants during their completion of the survey, we incorporated several attention checks throughout, as well as checks for access to both the auditory and visual information of the videos. Participants had to pass these checks in order to stay in the study. We are therefore confident that they understood and solved the tasks adequately. While we cannot rule out the possibility that some participants in online studies may not pay full attention to the tasks at hand, the same can be said for those who complete studies in laboratory settings as well. Further, it is likely that any such participants would be evenly distributed across study conditions.

Third, despite our use of a video that was more auditorily-engaging in Experiment 3, one could argue that both videos ultimately contained more visuospatial than auditory features, and thus would lead to a greater number of visuospatial cues available to trigger IEMs during the secondary task. As such, it could be argued that the results we obtained may not generalize to other videos, especially videos that may highlight auditory content to an even greater degree than in the present video (e.g., a video depicting a symphony performance, where less emphasis would be placed on the diversity of visuospatial elements). Obviously, in naturalistic contexts, the content of the involuntary memories depends on the characteristics of the original event. Here we controlled the modality of the cues and the sensory focus of participants' attention during encoding, and we included the same number of auditory versus visuospatial cues taken from the film, in order to control for such variation. Still, future research should look into the generalizability of the present findings in different contexts.

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research*. A definitive version was subsequently published in *Psychological Research.* DOI: 10.1007/s00426-020-01304-5.

Finally, despite our best attempts to equate the total number of visuospatial and auditory cues, it may still be the case that the visuospatial cues contained more details of the broader sensory environment from which they were taken than the auditory cues. As was mentioned above, this was due to limitations in our ability to completely isolate the visual fragments with our software capabilities. However, in this experiment we were primarily interested in the examination of the effect of visuospatial (as opposed to visual) cues on involuntary memories. In addition, the auditory cues used were also taken from the broader sensory environment in which they were embedded, providing a relative balance between the visuospatial and auditory cues. Of course, future research should consider examining the effect of visual and auditory cues alone (isolated from their broader sensory environment and spatial context) on involuntary memory elicitation and phenomenology.

In conclusion, the results of this series of experiments demonstrate the importance of the sensory modality of the retrieval cue in shaping the frequency and content of people's experience of involuntary episodic memories. The findings show that visuospatial cues trigger involuntary memories more frequently than auditory cues, especially of the visual parts of the events, suggesting that visuospatial cues may outshine auditory cues in terms of involuntary memory elicitation and the content of subsequent memories produced.

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research*. A definitive version was subsequently published in *Psychological Research.* DOI: 10.1007/s00426-020-01304-5.

Ethical Standards

This study was approved by the local ethics committee of the Department of Psychology and Behavioural Sciences at Aarhus University, and, as such, is in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments.

All participants gave their informed consent prior to their inclusion in the study.

The manuscript does not contain clinical studies or patient data.

References

Aydin, C. (2018). The differential contributions of visual imagery constructs on autobiographical thinking. *Memory, 26*, 189-200.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science, 4*, 417-423.

Berntsen, D. (1996). Involuntary autobiographical memories. *Applied Cognitive Psychology, 10*, 435–454.

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

Berntsen, D. (2009). *Involuntary autobiographical memories. An introduction to the unbidden past.* New York, NY: Cambridge University Press.

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

Berntsen, D., Rasmussen, A.S., Miles, A.N., Nielsen, N.P., & Ramsgaard, S.B. (2017). Spontaneous or intentional? Involuntary versus voluntary episodic memories in older and younger adults. *Psychology and Aging, 32*, 192-201.

Berntsen, D. & Rubin, D.C. (Eds.) (2012). *Understanding Autobiographical Memory: Theories and Approaches*. New York: Cambridge University Press.

Berntsen, D., Rubin, D.C., & Salgado, S. (2015). The frequency of involuntary autobiographical memories and future thoughts in relation to daydreaming, emotional distress, and age. *Consciousness and Cognition, 36*, 352-372.

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.

Buhrmester, M.D., Talaifar, S., & Gosling, S.D. (2018). An Evaluation of Amazon's Mechanical Turk, Its Rapid Rise, and Its Effective Use. *Perspectives on Psychological Science*, *13*, 149–154.

Congleton, A.R. & Berntsen, D. *Psychological Research* (2019). https://doi.org/10.1007/s00426-019-01215-0.

Ehlers, A. & Clark, D.M. (2000). A cognitive model of posttraumatic stress disorder. *Behaviour Research and Therapy, 38*, 319-345.

Ehlers, A., Hackmann, A., Steil, R., Clohessy, S., Wenninger, K., & Winter, H. (2002). The nature of intrusive memories after trauma: The warning signal hypothesis. *Behaviour Research and Therapy, 40*, 1021-1028.

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

Ehlers, A. & Steil, R. (1995). Maintenance of intrusive memories in posttraumatic stress disorder: A cognitive approach. *Behavioural and Cognitive Psychotherapy*, *23*, 217-249.

Grysman, A. (2014). The roles of gender and temporal distance in the recall of dissonant self-related memories. *Consciousness and Cognition*, *29*, 10-22.

Grysman, A., Prabhakar, J., Anglin, S.M., & Hudson, J.A. (2015). Self-enhancement and the life script in future thinking across the lifespan. *Memory*, *23*, 774-785.

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.

Horowitz, M.J. (1975). Intrusive and repetitive thought after experimental stress. *Archives of General Psychiatry, 32*, 1457-1463.

Horowitz, M.J. (1986). *Stress response syndromes* (2nd ed.). Northvale, NJ: Aronson.

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

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

Maillet, D. & Schacter, D.L. (2016). From mind wandering to involuntary retrieval: Age-related differences in spontaneous cognitive processes. *Neuropsychologia*, *80,* 142–156.

Mason, W. & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavioral Research Methods, 44*, 1-23.

Mazzoni, G., Vannucci, M., & Batool, I. (2014). Manipulating cues in involuntary autobiographical memory: Verbal cues are more effective than pictorial cues. *Memory & Cognition, 42*, 1076-1085.

McDaniel, M.A., Anderson, C.D., Einstein, G.O., & O'Halloran, C.M. (1988). Modulation of environmental reinstatement effects through encoding strategies. *American Journal of Psychology, 102*, 523-548.

Morris, C.D., Bransford, J.D., & Franks, J.J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning & Verbal Behavior, 16*, 519-533.

Rasmussen, A.S., & Berntsen, D. (2011). The unpredictable past: Spontaneous autobiographical memories outnumber autobiographical memories retrieved strategically. *Consciousness and Cognition, 20,* 1842–1846.

Rubin, D.C. & Berntsen, D. (2009). The Frequency of Voluntary and Involuntary Autobiographical Memories across the Lifespan. *Memory & Cognition, 37,* 679-688.

Rubin, D.C. & Umanath, S. (2015). Event memory: A theory of memory for laboratory, autobiographical and fictional events. *Psychological Review, 122*, 1-23.

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.

Schlagman, S., Kvavilashvili, L., & Schulz, J. (2007). Effects of age on involuntary autobiographical memories. In J. H. Mace (Ed.), *Involuntary memory* (pp. 87–112). Malden, MA: Blackwell.

Shipley, W.C. (1940). A self-administering scale for measuring intellectual impairment and deterioration. *The Journal of Psychology, 9*, 371-377.

Smith, S.M. (2013). Effects of environmental context on human memory. In T.J. Perfect and D.S. Lindsay (Eds.), *Sage handbook of applied memory* (pp. 162-182). London, UK: Sage Publications.

Smith, S.M. & Handy, J.D. (2014). Effects of varied and constant environmental contexts on acquisition and retention. *Journal of Experimental Psychology: Learning, Memory & Cognition, 40*, 1582-1593.

Smith, S.M., Handy, J.D., Hernandez, A., & Jacoby, L.L. (2018). Context specificity of automatic influences of memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 44*, 1501-1513.

Smith, S.M. & Manzano, I. (2010). Video context-dependent recall. *Behavior Research Methods, 42*, 292-301.

Tulving, E. (1984). Precis of elements of episodic memory. *Behavioral and Brain Sciences, 7,* 223–268.

Tulving, E. & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior, 5*, 381-391.

Watkins, O.C., & Watkins, M.J. (1975). Buildup of proactive inhibition as a cue-overload effect. *Journal of Experimental Psychology: Human Learning and Memory, 1*, 442–452.

Zachary, R.A. (1986). *Shipley Institute of Living Scale: Revised Manual*. Los Angeles: Western Psychological Services.

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research*. A definitive version was subsequently published in *Psychological Research.* DOI: 10.1007/s00426-020-01304-5.

Table 1

Experiment 1 means, (standard deviations), F-*values, and* η^2_p-*values of phenomenological questionnaire variables*

Variable	Visual Cues Only M SD	Auditory Cues Only M SD	*F*-values F(1,198)	η^2_p-values
Occurrence	6.04 (1.14)	5.38 (1.64)	10.91***	.05
Frequency ǂ	2.20 (0.44)	1.82 (0.63)	25.23***	.11
Same Scenes	3.78 (1.40)	3.75 (1.66)	.02	<.001
Visual Content	5.45 (1.40)	4.84 (1.71)	7.64**	.04
Auditory Content	3.11 (1.96)	4.31 (1.70)	21.37***	.10
Emotion	3.27 (1.67)	3.26 (1.72)	.002	<.001
Valence	3.56 (1.24)	3.53 (1.58)	.02	<.001

Note. * = p < .05, ** = p < .01, *** = p < .001, ǂ = these results are based on a square root transformation of the data

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research*. A definitive version was subsequently published in *Psychological Research.* DOI: 10.1007/s00426-020-01304-5.

Table 2

Experiment 2 means, (standard deviations), F*-values, and* η^2_p*-values of phenomenological questionnaire variables*

Variable	Visual Cues Only M SD	Auditory Cues Only M SD	Combined Cues M SD	*F*-values F(1,317)	η^2_p-values
Occurrence	6.05 (1.22)	5.48 (1.55)	5.81 (1.32)	4.38**	.03
Frequency ⸸	2.23 (0.56)	1.81 (0.62)	2.20 (0.49)	17.87***	.10
Same Scenes	3.78 (1.45)	3.59 (1.85)	3.91 (1.66)	1.01	.006
Visual Content	5.46 (1.26)	4.93 (1.67)	5.26 (1.43)	3.41*	.02
Auditory Content	2.93 (1.67)	4.37 (1.62)	3.64 (1.63)	19.22***	.11
Emotion	3.60 (1.62)	3.16 (1.53)	3.60 (1.63)	2.62	.02
Valence	3.84 (1.32)	3.76 (1.31)	3.45 (1.32)	2.79	.02

Note. * = p < .05, ** = p < .01, *** = p < .001, ⸸ = these results are based on a square root transformation of the data

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research*. A definitive version was subsequently published in *Psychological Research.* DOI: 10.1007/s00426-020-01304-5.

Table 3

Experiment 3 means, (standard deviations), F-*values, and* η^2_p-*values of phenomenological questionnaire variables*

Variable	Visual Cues Only		Auditory Cues Only		Combined Cues		*F*-values (η^2_p-values)		
	VisFix	AudFix	VisFix	AudFix	VisFix	AudFix	EnFix	Cue	Interaction
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	F(1,592)	F(2,592)	F(2,592)
Occurrence	6.35 (0.99)	6.25 (1.03)	6.05 (1.15)	6.10 (1.08)	6.37 (1.06)	6.34 (0.89)	.10 (.00)	4.04 (.013)*	.27 (.001)
Frequency ǂ	2.35 (0.41)	2.18 (0.52)	2.12 (0.41)	2.20 (0.40)	2.18 (0.41)	2.21 (0.44)	.26 (.00)	2.96 (.01)	4.52 (.02)*
Visual	5.68 (1.31)	5.47 (1.38)	5.56 (1.28)	5.70 (1.30)	5.54 (1.33)	5.62 (1.22)	.00 (.00)	.10 (.00)	1.08 (.004)
Auditory	4.37 (1.83)	4.58 (1.80)	4.56 (1.65)	4.86 (1.64)	4.61 (1.63)	4.97 (1.63)	4.42 (.01)*	1.83 (.01)	.09 (.00)

Note. VisFix = Visual Fixation, AudFix = Auditory Fixation, EnFix = Encoding Fixation, Cue = Cue Type, Visual = Visual Content, Auditory = Auditory Content, * = p < .05, ** = p < .01, *** = p < .001, ǂ = these results are based on a square root transformation of the data

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research*. A definitive version was subsequently published in *Psychological Research.* DOI: 10.1007/s00426-020-01304-5.

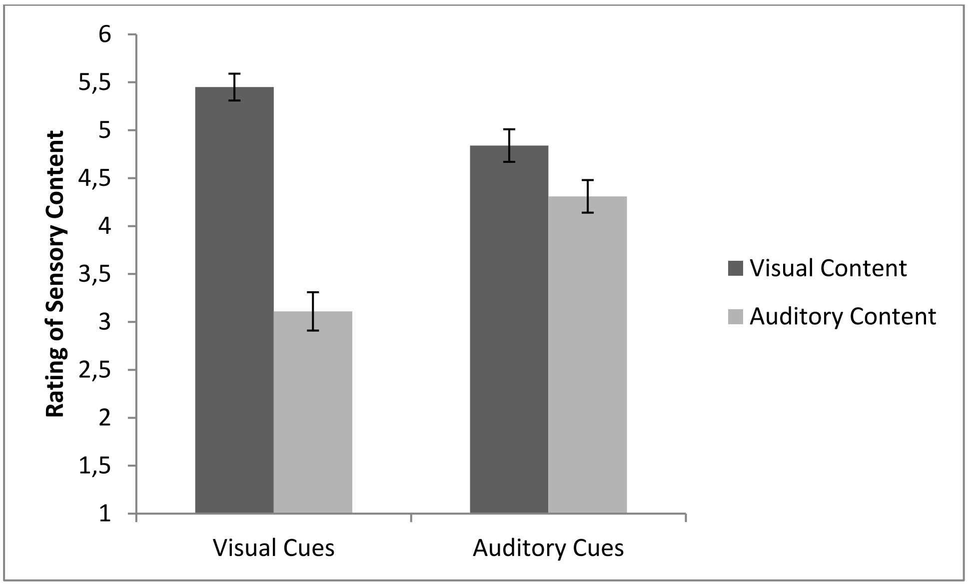

Figure 1. Experiment 1 interaction between Cue Type (Visual Cues vs. Auditory Cues) and Sensory Content (Visual Content vs. Auditory Content). Error bars represent standard error.

Notice: This is the author's version of a work that was accepted for publication in *Psychological Research*. A definitive version was subsequently published in *Psychological Research.* DOI: 10.1007/s00426-020-01304-5.

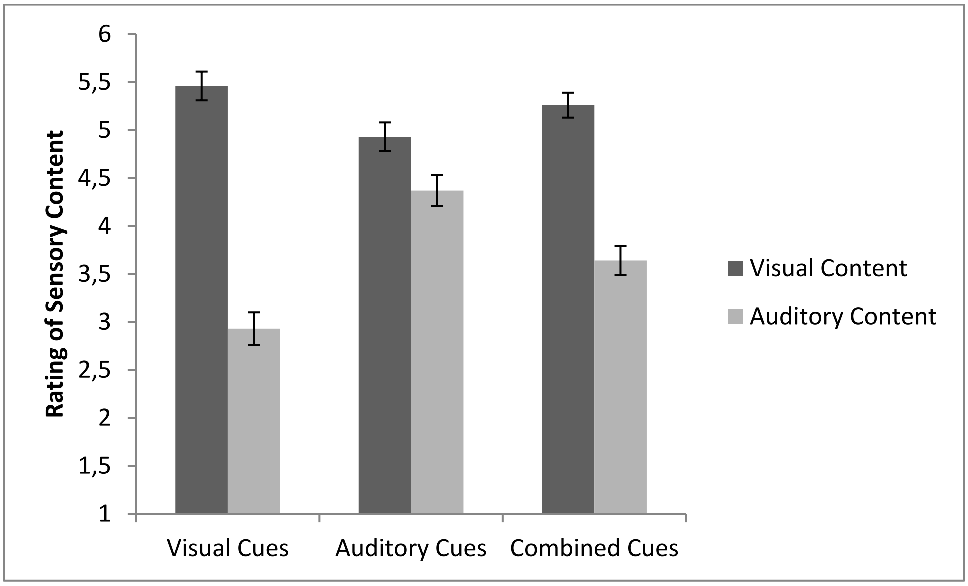

Figure 2. Experiment 2 interaction between Cue Type (Visual Cues vs. Auditory Cues vs. Combined Cues) and Sensory Content (Visual Content vs. Auditory Content). Error bars represent standard error.

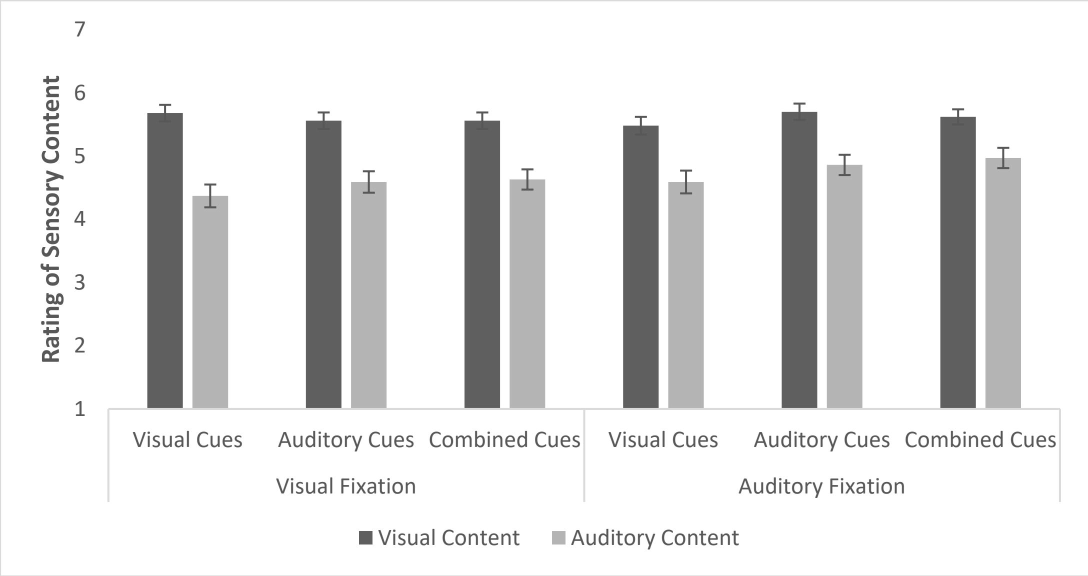

Figure 3. Experiment 3 interaction among Encoding Fixation (Visual Fixation vs. Auditory Fixation), Cue Type (Visual Cues vs. Auditory Cues vs. Combined Cues), and Sensory Content (Visual Content vs. Auditory Content). Error bars represent standard error.