Effects of Scene Memorability on Eyewitness Memory

Bennet Hieronymi

S4291980

Department of Psychology, University of Groningen

PSB3E-BTHO.2022-2023.1: Bachelor Honours Thesis

Supervisor: Dr. Mark Nieuwenstein

July 24, 2023

Abstract

Eyewitness testimony is subject to bias and environmental influences and can be the source of wrongful convictions. This study investigates whether a memorable crime scene will lead to stronger episodic memory associations between the elements that make up a crime event, thereby benefiting cued recognition of the culprit in a lineup.

Analogous to the formation of the episodic memory, our experiment consists of two phases. The first phase, in which the participants (N = 113) were shown three different images, is called the encoding phase. The images depicted either a crime scene, a culprit or a stolen object. While the pictures of the culprit and the stolen object remained unchanged for all participants, the presented crime scenes differed in their memorability, which was objectified by means of a memorability score. The intervention groups were thus divided into memorable and non-memorable according to the crime scene conditions. In the subsequent second phase, the participants equally completed a cued retrieval task. Participants were asked to identify the culprit in a four-options forced choice format, using either the scene image or the object image as a cue. Results show that participants identified the culprit more often in the memorable conditions. Therefore, we conclude that scene memorability positively influences identification accuracy and thus increases the reliability of eyewitness testimony.

Additionally, this experimental design provided data about the associations in an episodic memory. We asked whether the above-stated positive influence of scene memorability translates to all associations in an episodic memory. We compared the memorability effect across the different cueing conditions, thereby analysing the difference in associations of scene and culprit and respectively object and culprit. The memorability effect was detected in both cueing conditions. Thus, the strength of one association in episodic memory was correlated to the strength of other associations in the same episodic memory. This indicates dependence between associations retrieved from episodic memory.

Keywords: Eyewitness Memory, Episodic Memory, Memorability, Dependent Theory

Effects of Scene Memorability on Eyewitness Memory

The forensic psychology community has developed a thorough understanding of memory evidence as a type of 'trace evidence', characterised by the difficulty of collection and ease of contamination (Doyle, 2014). Eyewitness memory does not function like a videotape recording that captures every detail and allows for accurate retrieval. On the contrary, it is generally accepted that memory formation allows for contamination at all levels of information processing (Doyle, 2014). In a landmark study by Loftus et al., (1978), participants who were shown a “stop” sign later falsely reported seeing a "give way" sign when they had also been shown a "give way" sign. This indicates that eyewitness memory is malleable.

Parameters that influence the reliability of eyewitness memory can be divided into two categories: characteristics of the observed event and characteristics of the witness (Pezdek, 2012). Characteristics of the witness include witness confidence (Sporer et al., 1995), cross-racial identification (Meisner & Brigham, 2001), eyewitness stress (Deffenbacher et al., 2004), and witness intoxication (Dysart et al., 2002). In practice, characteristics of the observed event have been shown to have a greater effect on eyewitness memory than characteristics of the witness (Pezdek, 2012). This can be demonstrated by the following example. Suppose a witness observes a crime in poor lighting conditions and is asked to recall information months after the event. Even if the witness had unimpaired visual memory, the reliability of the eyewitness memory would be affected by the lighting conditions of the crime scene. Thus, the characteristics of the observed event outweigh the characteristics of the witness (Pezdek, 2012). One specific characteristic of the crime event is the memorability of the scene, which will be discussed in detail in the following chapter. Although widely studied, so far memorability has not been introduced into the field of eyewitness research.

Memorability

Isola et al. (2011) have defined memorability as "the probability that an observer will detect a repetition of a photograph a few minutes after exposure when it is presented in a stream of images". Humans have a great capacity to remember images even after a short exposure. Skimming through a magazine, for example, can be enough to memorise certain images. Khosla et al. (2015) have observed a high degree of consistency between people in which images are remembered, despite large individual or contextual differences. Image properties such as the information displayed, visual salience, and imageability can only partially explain or predict whether an image will be remembered (Isola et al., 2013). We, therefore, assume that images have an innate property that partially determines the memorability of those very images. Studies further show that the memorability of an image is stable over time (Borkin et al., 2013).

Effect of Memorability on Associative Memory

In psychology, associative memory describes the ability to learn and remember the relationship between unrelated elements of a memory. The formation of memories of two paired stimuli leads to pairwise associations between these stimuli. The associative nature of memories is commonly tested using a cued retrieval task. A retrieval cue is presented to evoke the memory of a retrieval target (Moult, 2011). The imageability of a stimulus positively influences performance in a cued retrieval task. This can be explained by two different effects. Firstly, the high retrievability of the retrieval target makes it easier to remember. Secondly, the imageability of one stimulus strengthens the association between the retrieval cue and the retrieval target which leads to enhanced retrieval performance. The strength of associations between memories is in turn influenced by imageability (Madan et al., 2010). In their study, Madan et al. (2010) presented pairs of words (A and B) to the participants varying in imageability (high vs. low). Pairs were either “pure” (high – high, low – low) or “mixed” (high – low, low – high). Subsequently, participants performed a cued retrieval task, either receiving the cue A (forward) or the cue B (backward). A crucial assumption was that if the retrievability of the target item caused the positive effect of imageability on retrieval performance, the results would show a significant difference in the comparison between forward and backward cueing in the mixed pair conditions. Higher retrieval performance would be expected when the highly imageable stimuli are the retrieval targets and lower retrieval performance when the highly imageable stimuli are the retrieval cues. Analysis of the data did not reveal a significant difference in retrieval performance in the mixed pair condition, suggesting that the effect of imageability on retrieval performance was due to its influence on the strength of the association between the stimuli. Therefore, Madan et al. (2010) conclude that variation in imageability influences the strength of the associations between items. We aim to extend their findings by including the memorability of a stimulus, which is partly determined by its imageability. We also ask whether high memorability extends to all associations in a complex multimodal memory consisting of three instead of two stimuli. By manipulating the memorability of stimuli that could be part of a crime scene, we will link this research objective to the study of eyewitness memory.

Neurological research highlights the influence of memorability on associative memory performance. Xie et al (2020) found that a memorable element of an episodic memory increases retrieval rates, even when associated with a non-memorable cue. Their results can be explained by the assumption that memorable elements are more easily retrieved than non-memorable elements. When a cue is presented, a cognitive search process is initiated to retrieve the missing components of the encoded association. Rather than searching the brain at random, the search process begins in a specific region of the brain where highly memorable stimuli are stored (Xie et al., 2020). This neurological pathway leads to a rapid cognitive link between recurring, highly memorable stimuli and their associations (Xie et al., 2020).

Eyewitness Memory Relies on Episodic Memory

There are two different concepts of memory: Episodic memory and semantic memory. Semantic memory holds conceptual and factual knowledge (Tulving, 1985). In contrast, episodic memory is characterised by the vivid and detailed retrieval of evocative memories from the past (Tulving, 1984). Episodic memory allows the conscious recollection of personal experiences, including a subjective perception of oneself, the event, and the place and time of the event. Eyewitness memory is therefore based on episodic memory encoded for the crime observed.

Typically, eyewitnesses perceive themselves as part of the crime scene and remember individual elements of the crime. The different elements encoded in the episodic memory are represented in different neocortical areas of the brain and processed into a coherent memory (Horner et al., 2015). It is thought that the hippocampus mediates associations between individual elements of an episodic memory. The hippocampal networks bind multimodal elements of an experienced event into a single engram (Horner & Burgess, 2013). Tulving (1985) first defined these complex multimodal events as engrams. Because the elements of an engram are stored separately, but the experienced event is retrieved as a whole, Horner and Burgess (2013) predict a relationship between the elements. They assume that the elements of an engram are stored as multiple pairwise associations between individual elements comprising the engram. Piecing together fragmented information to recall an event is a form of holistic processing. The idea was introduced by Jones (1976) in his “Fragmentation Hypothesis“. Jones (1976) referred to engrams as memory fragments and stated that when one part of a fragment is used as a cue, it can serve as a gateway to access all other elements of that fragment. Holistic encoding is also a key concept in the “Dependent Theory” by

Horner and Burgess (2013). The “Dependent Theory” states that all event elements are associated and that the strength of an association such as A and B correlates to the strengths of another association C and B, suggesting dependence between associations of an engram.

Eyewitness memory also consists of different elements which can act as a cue to retrieve associated elements of the same memory. For example, an eyewitness to a theft may remember different features of the crime scene, possibly the scenery, the culprit, and the stolen object. The presentation of one element of a memory acts as a cue, increasing the likelihood of retrieval of other elements of the same memory (Meiser & Bröder, 2002). This has direct practical relevance. Images of the crime scene and the stolen object are usually available during interrogations and could therefore act as a cue.

In a hypothetical engram consisting of three elements A, B, and C, the cueing of element A would lead to the activation of the associations A-B and A-C. Thus, elements B and C are retrieved. If element B is the retrieval target, it would be superfluous to cue A and C because cueing A would already retrieve C. Thus, using multiple elements of a fragment as cues is unnecessary and does not significantly improve recall performance.

Research Paradigm

There are certain requirements that an episodic memory research paradigm must meet (Pause et al., 2013). First, participants in an experiment should not be instructed to remember the stimuli presented. Forming a memory with the aim of remembering activates the semantic memory system. For the formation of episodic memory, it is crucial to imagine oneself as part of the spatial and temporal context of the event, rather than trying to remember the presented elements. Second, episodic memory is often a representation of an important life event that is often accompanied by, and probably induced by, emotional activation. Therefore, the induction of an episodic memory requires emotional valence. Third, the experimental design should involve only a single presentation of the episodic memory items, since the formation of episodic memory is by definition a one-trial learning event. Repeated presentation of the items would lead to a mediating role of the semantic memory system. Fourth, the episodic information induced and retrieved should include information about what happened, where it happened, and when it happened. It is not crucial to include a time stamp, but rather that the person remembering the event knows when to place this memory in the timeline of his or her life. Other episodic memories can serve as a point of reference, e.g. I witnessed this crime shortly after my parents separated.

After establishing the standards of episodic memory research, the second step was to check whether these were met in eyewitness research. A common method in eyewitness research is the presentation of videos of the crime scene to elicit an episodic memory of the crime event (List, 1986; Wulff & Hyman Jr., 2020). However, instructing participants to recall the presented event after the video presentation can activate the semantic memory system (Pause et al., 2013). The interview format itself can also introduce bias if questions are framed in a leading way or lack ecological validity. Ecological validity is particularly important when transferring from scientific simulation to real witness interviews.

We decided to implement our research question in the well-established episodic memory research paradigm of Horner and Burgess (2013). Their episodic memory research paradigm consists of an encoding phase and a cued retrieval phase (Horner & Burgess, 2013). A crucial element of this experiment is to instruct participants to vividly imagine that the three elements presented during the encoding phase are interacting with each other. This allows for the formation of experimentally induced episodic memory. In our experiment, participants are asked to vividly imagine being at a crime scene and that the culprit is stealing an object. This experimental set-up includes all aspects of an episodic memory, i.e., spatial and personal information as well as an emotion-evoking scenario (Pause et al. 2013). In the cued retrieval task, participants must identify the culprit cued by either the stolen object or the scene. The use of a set-up consisting of the three elements (triplet) including culprit, stolen object and scene, rather than a set-up consisting of two elements (doublet), is crucial for assessing dependence between the associations of an engram. In line with the "Dependent Theory" proposed by Horner and Burgess (2013), we assume a dependency between the associations of an episodic memory.

The research paradigm by Horner and Burgess (2013) was previously adapted for eyewitness memory research by Erickson et al. (2022). In their study, they investigated the role of weapon involvement and racial biases on eyewitness memory. They also used this approach to assess the age-related deficit in associative memory in the context of eyewitness testimony. Their research demonstrates the advantages of applying an episodic memory research paradigm to eyewitness memory research. Firstly, the paradigm allows the induction of episodic memory and its implementation in an experimental setting. The formation of an episodic memory creates an ecologically reliable experiment. Secondly, the paradigm also accounts for the multiple association between individual elements of an engram. Thus, Erickson et al. (2022) argue that the episodic memory research paradigm is applicable to both forensic and cognitive theories. We developed such hypotheses, which we analysed sequentially from the same data set.

Firstly, we will analyse whether the memorability of the scene image influences the retrieval accuracy during cued retrieval. We hypothesise greater recall accuracy for triplets including a highly memorable scene image. We base our hypothesis on the research findings of Madan et al. (2010), who propose that the imageability, a determinant of the memorability, of a word influences the strength of pairwise associations between that word and other elements associated with that word during encoding.

Secondly, we hypothesise dependency between associations of an engram. We aim to reproduce the “Dependent Theory” introduced by Horner and Burgess (2013). We will analyse the correlation between memorability of the presented crime scene and the recognition rate by the participant. We account for the influencing variables of the different cues, namely crime object or crime scene. We argue that detecting a positive memorability effect when given the scene as a cue and given the object as a cue implies a pairwise association between the scene and culprit respectively the pairwise association between object and culprit. From this we infer a strengthened association between the two cues, scene and object, and the culprit. Thus, the memorability effect is transferred from one association of the triplet to each of the remaining associations. This indicates dependence between the associations of an engram.

Difference to the Pilot Study

The current study is a pre-registered confirmatory replication of a previous pilot study. In pilot study, a total of 56 participants were presented with sets of three pictures depicting a crime scene, a culprit, and a stolen object. Participants were asked to imagine themselves as part of a crime scene that included all three images. Half of the image triplets contained a highly memorable image of the crime scene. In the later cued recall, we assessed the participants' ability to recognise the different elements of the crime events. Participants received a cue of one or two elements of the presented triplet. The pilot study consisted of three experimental conditions. The experimental conditions differed in terms of the cue presented. In the first condition, the image of the crime scene was presented as a cue and participants were asked to identify the associated culprit. The second condition included all elements of the presented triplet, with the scene and object as cues and the image of the culprit as the recall target. The third condition consisted of cueing the culprit and targeting the object. The results showed that participants were more accurate at recognising elements from highly memorable pairings. The comparison of the different conditions provided modest support for the "Dependent Theory" by Horner and Burgess (2013). However, it is important to note that the small sample size of the study limited its statistical power. Based on the results of the pilot study, we decided to make three methodological changes.

First, to avoid the influence of coincidental differences in memory for the object-face pairs, the current study will counterbalance the assignment of object-face pairs to the memorable and non-memorable scenes. This was not the case in the pilot study, where the object-face pairs were assigned to a scene image that was either memorable or non-memorable.

Second, we increased the number of encoding trials from 20 to 28. Thus, we included 14 indoor scenes instead of ten. By increasing the number of trials per condition, we aim to increase the correlation between repeated measures, and, simultaneously, the effect size (Brysbaert, 2019). Our decision to include 14 indoor scenes was based on pilot data showing a linear increase in repeated-measures correlation (i.e., average percent correct culprit identification for memorable and non-memorable scenes) when subsampling from 4 to 10 observations. When we subsampled the scene trials from the pilot study, we found a linear increase in the correlation. Taking advantage of this increase, we found that increasing the number of scenes from 10 to 14 would be expected to increase the correlation from 0.62 to 0.79. If we assume that the current study would produce comparable means and standard deviations as the pilot study, the expected increase in correlation would result in an expected increase in effect size from .25 to .38. With the increase in encoding trials, we decided to present the triplets in four blocks of seven, whereas the pilot study presented the triplets in two blocks of ten.

The third methodological difference from the pilot study is the cue type conditions and the within-subject design. This study focuses on the identification of the culprit cueing the scene and object. Each participant will identify a culprit twice, once with the scene cue and once with the object cue. We have decided to deviate from the pilot study design in this respect because retrieving the culprit has more ecological validity than identifying the crime scene or the stolen object. In a real-life theft scenario, these factors are usually known, whereas the identification of the culprit is crucial. Furthermore, we are interested in whether the strength of the association between the scene and the culprit differs from the strength of the association between the object and the culprit. By comparing the interaction effect between the two cueing conditions and memorability, we can analyse whether the effect of the memorable stimuli (scene) transfers equally to all associations within an engram (Horner & Burgess, 2013; Horner et al., 2014).

Methods

Participants

We gathered data through Prolific (prolific.co). Participants were eligible to participate if they were between 18 and 40 years old and English-speaking. We derived a desired sample size of 120 in consideration of the pilot data which showed a small effect size of dz = .247 for the main effect of memorability. A power analysis using G-Power (Faul et al., 2007) identified that a minimum sample of 103 participants is required to have 80% power for detecting this effect with alpha = .05. In addition, we predicted that the effect of memorability will be equally strong for the two types of cued retrieval tasks. In order to detect moderate evidence for this null hypothesis in a Bayesian analysis we would need 60 participants (Brysbaert, 2019, p. 9). Therefore, the intended sample of 120 participants would be sufficient to detect a small effect for memorability and at least moderate evidence for the null hypothesis that this effect would not be different depending on whether the scene image or the object image was used as a cue.

The final sample consisted of 113 participants with 58 male and 55 female. The mean age was 26.4 years. The loss of data for 7 participants was due to technical issues and the fact that one participant revoked consent. The study employed specific exclusion criteria to ensure the quality and reliability of the data. Participants who performed at chance level (0.25) during the cued retrieval task would have been excluded from the analysis. Within the sample, no participant performed at chance.

Materials

The experiment was programmed using the software OpenSesame (Mathôt et al., 2012). The experiment was implemented into online Jatos studies with a display resolution of 1024x768 pixels. Participants were instructed to be in a quiet environment to avoid distraction. Additionally, participation was only possible while using a desktop computer or laptop.

Faces*:*

The Face pictures used in the study were selected from the 10K Adult Faces Database (Bainbridge et al., 2013). Within this database, two thousand face images were rated for various factors related to facial attributes. Furthermore, the database provides the memorability score for each face image. During the selection of images, we controlled for age and gender, including images of faces in the age range of 30 to 40 and an equal number of males and females. Furthermore, we excluded faces with special appearances such as face tattoos. The selected pictures had a hit rate between .3 and .4, which indicates low memorability. Only the faces of people judged to be of white ethnicity were included in the selection to prevent any potential influences of racial biases in eyewitness testimony (Meissner & Brigham, 2001). Erickson et al. (2022) suggested that suspects who appear angry are more likely to be identified as the culprit. Therefore, we selected face images scoring four or lower on a nine-point Likert scale on the facial attributes of "aggressiveness", "unfriendliness", and "untrustworthiness".

Objects:

We derived the images of objects from the Things database (Hebart et al., 2019). Kramer et al. (2022) provided memorability scores for the images in this database. We selected a total of 28 images. Objects were chosen accordingly and semantically matched with the scenes to enable a realistic crime scenario (see Appendix A for Table A1). The memorability score of the images ranged from .7 to .8.

Scenes:

We selected two images for each of the 14 categories of indoor scenes (e.g., supermarket, office, attic, etc.). One image had a high-memorability score and the other had a low memorability score. We selected these images from the image memorability database of Isola et al. (2011). The hit-rate ranges and averages for the non-memorable and memorable images were .46- .62 (M = .56) and .83- .93 (M = .87), respectively.

Design and Procedure

In order to test the above-mentioned hypotheses, we adapted a research paradigm from Horner and Burgess (2013). The research paradigm consists of an encoding phase and a cued retrieval task. The study has a within-subject design with two independent variables, namely scene memorability (memorable vs. non-memorable) and cue type (object vs. cue). The dependent variable is the correct percentage for cued recognition of the culprit. During the encoding phase, we present participants 28 encoding trials consisting of a triplet of images displaying a picture of an indoor scene, a face, and an object. For half of the 28 encoding trials the indoor scene images have a high memorability score and for the other half a low memorability score (e.g., a memorable and non-memorable image of an attic). Each triplet of images is presented for 20 seconds, allowing participants to imagine a vivid scenario of the crime event. The assignment of face-object pairs to memorable and non-memorable scene images in the triplets are counterbalanced across participants. Participants are instructed to imagine themselves standing in a scenery while the person whose face is depicted steals the object. In doing so, participants are asked to be as imaginative as possible to create a vivid scenario of the crime event. This allows the formation of an episodic memory. The encoding trials are presented in a block-style fashion. Within a block seven triplets are presented either with 3 memorable (or four memorable) scene images and 4 non-memorable (3 non-memorable) scene images. Thus, we split the presentation of encoding trials into four blocks. In between blocks, participants have the opportunity to take a short break. The order of the encoding trials with different crime scenes is randomized.

Following the encoding phase, participants performed a cued recognition task in which they were presented with a cue that is categorised as an object cue or a scene cue, and a series of four face images (the target culprit and three other culprits of the same attributed gender) from which they had to identify the culprit. Each culprit image is assigned a number from one to four. Participants must indicate which culprit is associated with the presented cue by pressing the corresponding key (1, 2, 3, or 4; see Figure A1). Participants are asked to identify the culprit twice. Once with the 'scene cue' and once with the 'object cue'. Participants were randomised as to which cued retrieval task they performed first. Approximately half of the participants performed the cued retrieval task first, receiving the object cue and then the scene cue.

Results

Remark Results

There was an error in the lineup of faces for one of the trials, such that the face of the culprit for one of the scenarios was presented twice in the lineup. Therefore, the recognition task had two correct answers. Importantly, however, this error occurred only once in each memorability condition and it thus did not influence the comparison of accuracy for trials with memorable and non-memorable scenes. We coded any selection of the correct culprit as a correct response for these trials.

Preliminary Analysis: Assumption Check

The analysis of our data was done in a 2x2 repeated measure ANOVA with two within-subject factors (scene memorability and cue type) both with two levels. We computed the average proportion correct for each combination of scene memorability and cue type. Due to the large sample size of N=113, we assume the normality assumption to be met (Ghasemi & Zahediasl, 2012). In a confirmatory normality test the assumption was met with a Shapiro-Wilk test (W= 0.989, p = .536). Furthermore, this analysis has two levels and therefore no violation of the sphericity assumption was possible.

The Influence of Scene Memorability on Eyewitness Identification

The 2x2 repeated measures ANOVA revealed a non-significant main effect of scene memorability on culprit identification accuracy (p = .060), F(1, 112) = 3,606, partial η^2 = 0.31. However, since we hypothesised better performance for the memorable scene condition we conducted a one-sided paired sample t-test. There is a significant difference in the identification accuracy for the memorable scene condition (M = .621, SD = .216) and non-memorable scene condition (M = .597, SD = .224); t(112) = 1.899, p = .030. An estimate of the effect size with Cohen's d of d=.180 is seen as very small. These results indicate greater accuracy of culprit identification of triplets in the memorable scene condition (see Appendix B for Figure B1). We tested the same hypothesis with a more conservative Bayesian one-tailed t-test. This analysis resulted in a Bayes' Factor of BF10 = 1.350 indicating anecdotal evidence for the presence of an effect of scene memorability.

Dependency Analysis of the Associations of an Episodic Memory

The second hypothesis was that the effect of the scene memorability would not differ between cue-type conditions. This is equivalent to testing the significance of the interaction effect between memorability and cue type in the 2x2 repeated measures ANOVA. The statistical analysis cannot detect a significant interaction effect between memorability and cue type (p = 0.952), $F(1,112) = 0.004$, partial $\eta^2 = 0.00$. This supports our second hypothesis and is in line with the “Dependent Theory“ by Horner and Burgess (2013). We tested the same hypothesis with a more conservative Bayesian repeated measures ANOVA. This analysis resulted in a Bayes’ Factor of BF10 = 0.652 indicating strong evidence for the null hypothesis of the interaction effect.

Effect of Cue Type on Eyewitness Identification Accuracy

There is a significant effect of cue type on culprit identification accuracy (p = 0.023). $F(1, 112) = 5.306$, partial $\eta^2 = 0.45$. This indicates a significant difference in retrieval accuracy between both cueing conditions (object vs. scene). The data suggests that the object images functioned as a better cue than scene images.

Robustness Analysis of the Main Effect of Memorability

We conducted further analysis to explore the robustness of the above-stated main effect of memorability. More specifically, we looked at the memorability effect across participants, scene types, and presentation blocks. We aim to evaluate the consistency of the observed memorability benefit per participant, for the different scenes used, and for the four presentation blocks.

First, we calculated a difference score between the mean recognition accuracy of the memorable scene condition and the non-memorable scene condition for each participant. The calculated value will be referred to as the memorability benefit score in the following passages. Out of the total 113 participants, 59 participants obtained a positive score, indicating a higher recognition rate for the memorable scene condition. Conversely, 41 participants received a negative score, indicating a lower recognition rate for the memorable scene condition. The remaining 13 participants obtained a null result, indicating no detectable difference in recognition accuracy under memorable respectively non-memorable conditions (see Figure B2).

We also examined the consistency of the main effect of memorability for each scene type. The influence of memorability was again assessed by calculating a difference score for recognition accuracy between the two levels (memorable vs. non-memorable). We calculated values for all individual scene types used in the experimental setup to determine the consistency of the memorability benefit. Consecutive analysis of the difference scores shows eight of the scenes yielding a positive score, indicating a higher recognition rate in the memorable scene condition. On the opposite, six scenes produced a negative score, indicating that recognition accuracy was higher for the non-memorable scene triplets (see Figure B3).

As described in the Methods section, the encoding phases consisted of four blocks. We. analysed the consistency of the memorability effect across presentation blocks. To do this, we calculated the difference (memorability benefit score) between the average accuracy of the memorable and non-memorable scene conditions for each presentation block. In presentation blocks two, three and four, the memorable scene condition shows higher identification accuracy than the non-memorable scene condition. In contrast, in presentation block one there is a small difference between the memorability conditions, with greater identification accuracy for the non-memorable scene condition (see Figure B4). In addition, we conducted a repeated measures ANOVA to identify whether there is a significant difference between the culprit identification accuracy across presentation blocks. There was a significant difference in the culprit identification accuracy in at least two presentation blocks ($p < 0.001$), $F(3, 336) = 9{,}779$, partial $\eta^2 = 0.63$ (see Figure B5). A visual representation of the mean culprit identification per presentation block indicates on average a greater accuracy in the first and last presentation blocks.

Analysis of Response Times in the Cued Recognition Task

Our analyses showed that the effect of scene memorability did not differ between cue types, suggesting that scene memorability had a similar influence on scene-culprit associations as it did on object-culprit associations. Although the results indicate a dependency between the individual associations, they do not exclude the possibility that only the associations between the memorable stimulus (scene image) and the directly associated elements are strengthened. This would lead to enhanced associations between scene and culprit and scene and object. The association between the object and the culprit would remain unaffected. This would still lead to the same memorability effects in both cueing conditions when participants retrieve the information about the culprit sequentially. For example, a participant receives an object as a cue and retrieves the associated culprit through the scene. In order to test this alternative interpretation, we examined participants' reaction times. The idea is that if the participants were to retrieve the culprit sequentially, the reaction times for the object cueing condition would increase. We only included trials in which the culprit was correctly identified. In contrast to our expectation, we observed that participants needed significantly more time to identify the culprit when receiving the scene image as a cue. The cue type had a significant effect on reaction time ($p = 0.002$), $F(1, 112) = 9{,}779$, partial $\eta^2 = 0.08$. The effect size estimate calculated as partial eta squared ($\eta^2 = 0.08$) indicates a medium effect.

Discussion

Over the past four decades, eyewitness memory research has highlighted the limitations of eyewitness testimony. The increased research effort has led to a shift in public perception towards a more critical assessment of the reliability of eyewitness testimony and has contributed to the implementation of forensic psychological theories in the courts (Doyle, 2014). However, it remains challenging to objectively distinguish between reliable and unreliable eyewitness memory. The aim of this study was to assess the effect of the memorability of a crime scene on the reliability of eyewitness identification. We implemented this research question in a well-established episodic memory research paradigm by Horner and Burgess (2013; Horner et al., 2014). In addition, this research paradigm allowed for the analysis of engram retrieval of episodic memories.

The combination of the episodic memory research paradigm and the effect of memorability on eyewitness memory is a novel approach. Erickson et al. (2022) describe the application of episodic memory research paradigms to eyewitness memory research as promising due to its increased ecological validity and ease of implementation in the experimental setting.

The data collected supports our first hypothesis about the benefit of memorable crime scenes on culprit identification. We found a positive effect of memorable scenes on culprit identification by the eyewitness. Thus, triplets with a memorable scene image had significantly better culprit identification rates. In an exploratory robustness analysis, we detect the memorability benefit for the majority of participants, scene types, and presentation blocks. The robustness analysis underlines the consistency of the memorability effect.

Secondly, we aimed to reproduce the "Dependent Theory" by Horner and Burgess (2013), showing that the strength of one association of an engram depends on the strength of other association in the same engram. To analyse the dependency between associations we compared the memorability influence between different associations in an engram. We found no significant interaction between memorability and cue type, indicating that the effect of memorability was equally in both cueing conditions. Thus, the memorability effect is transferred from one association of the triplet to each of the remaining associations. This indicates dependence between the associations of an engram.

Within trials of the object cue type condition, participants identified the culprits faster and more accurately than in the scene cue type condition. The theoretical implications are discussed below.

The Influence of Scene Memorability on Eyewitness Identification

As hypothesised, the main effect of memorability was significant, indicating better recognition performance for the memorable scene condition. We conclude that the memorability of the crime scene image influences the reliability of eyewitness testimony. This is in line with previous research by Madan et al. (2010), who described that the imageability of a stimulus can strengthen associations between elements of the same engram. Thus, we proposed that the high memorability of the scenes, which is partially determined by its imageability, strengthened the associations between scene, object and culprit. This effect was reflected in greater retrieval accuracy for triplets associated with memorable scenes. The data underline that in both cueing trials, object and scene, participants had better culprit identification accuracy for memorable scene triplets (see Figure n.). However, the detected effect size found (d = 0.18) is considered very small and leads to limited practical implications of our findings. In a separate analysis of the data, we tested the robustness of the memorability effect. The data suggest that most participants showed a memorability advantage and that it was equally present across most scene types and presentation blocks.

Interaction Effect Memorability and Cue Type

The interaction effect between memorability and cue type was not significant, indicating that the memorability effect did not differ between the two cue type conditions. We hypothesised a null effect of the memorability effect between cue conditions, which is supported by the data.

The memorability of an image strengthens the pairwise associations between elements of an engram (Madan et al., 2010). In consideration of Jones (1976) and Horner and Burgess (2013), we assumed associations between all elements of an episodic memory. Thus, the memorability benefit of the scene would be transferred to all associations of an engram. Therefore, we tested the dependency between associations by analysing the memorability benefit between cueing conditions.

If the associations of an episodic memory are not dependent on each other (independent encoding), the data should have shown a significant difference in the memorability effect between cueing conditions. We would have expected a larger memorability effect for triplets of the memorable scene condition. The data suggest that the memorability effect was also found for the object cueing condition, which is inconsistent with independent encoding. It appears that the memorability effect of the memorable scene translates to the pairwise associations between all elements of the triplet (Madan et al., 2010). Therefore, our data support the “Dependent Theory” (Horner & Burgess, 2013).

Madan et al. (2010) detected a positive influence of imageability on the association between two elements. Therefore, only It could be argued that the memorability of the scene image only strengthens the pairwise association between scene and face and scene and object. Therefore, the pairwise association between object and face would remain normal. The results showed that the memorability effect was present for both cue types. This could be explained by sequential retrieval processes, where participants receive the object cue and retrieve the culprit through the memorable scene. For example, a participant is cued with an image of a blender (associated with a culprit and the memorable kitchen) and recalls the memorable image of the kitchen and then the culprit associated with the memorable kitchen. This possible sequential retrieval process offers an alternative interpretation of our explanation of association dependency. Sequential retrieval consists of the activation of multiple associations and item retrieval (object to scene to culprit). Therefore, we hypothesised that sequential retrieval would require more time to identify the culprit than direct retrieval (scene to culprit). We analysed the reaction time required by each participant to correctly identify the culprit and compared it between cue types. The results showed a significant difference in reaction time between the two cueing conditions. Participants needed significantly more time to correctly identify the culprit when they received the scene image as a cue. This contradicts our hypothesis of sequential retrieval of the culprit.

Explanatory Discussion of the Cue Type Effect

The results show a significant difference between the cueing conditions for identification accuracy and reaction time. Within the cue-object condition, participants had better culprit identification accuracy and shorter reaction time. Reaction time is commonly used as an indication of cognitive effort (Robinson et al., 1997). Considering the above results, the object image seems to be the better cue. In evaluating these findings, we found a possible explanation for the cue type effect in human information processing. Adults have the ability to make connections between elements such as objects, events, or people (Horn et al., 2021). In memory tasks involving semantically related items, adults group information by category and recall related items adjacently. This phenomenon is known as category clustering (Horn et al., 2021).

Participants may have stored semantically related information, such as different images of the same scene, in one category. For example, the participant might imagine a theft scenario in a memorable kitchen and later imagines another crime event in a non-memorable kitchen. Both scenarios could be encoded in the category kitchen. Potentially, during the retrieval task, the non-memorable kitchen scene image appears, and the participant could confuse information associated with the memorable kitchen crime event. This processing is represented in the pair-clustering multinomial model (Batchelder &Riefer, 1980, 1986). This model summarises the potential processing of associated information (e.g. Kitchen A and Kitchen B, see Appendix C for model visualisation). The model describes the probability of clustered information processing with parameter c, the probability of clustered retrieval (with probability r), and the probability of non-clustered retrieval (with probability 1 - r). Clustered processing (with probability c) would lead to grouped encoding of semantically related information such as Kitchen A and Kitchen B. With probability r, both related elements are retrieved together. The retrieval of both elements could lead to the retrieval of their associated crime scene elements ("culprit A" and "object A" and "culprit B" and "object B"). Therefore, cueing an image of Kitchen A or B could lead to the retrieval of multiple engrams. Processing multiple engrams could ultimately lead to longer reaction times and higher error rates in identifying the culprit. Category clustering is less likely to happen between the object stimuli, as the objects associated with one scene type were only vaguely semantically related to each other (see Table A1). Category clustering could explain the different reaction time and retrieval accuracy between cue type conditions (object vs. scene).

Implications

Although a large body of research has been created to identify and administer the fact that eyewitness memory is malleable and prone to contamination, we still cannot say with certainty how we distinguish between reliable and unreliable eyewitness testimony. This study helped to shed light on the influence of the memorability of the crime scene on the reliability of eyewitness memory. Furthermore, the processing of information such as episodic memory engrams is crucial to further understanding cognitive processes. In this study, we replicated the findings of Horner and Burgess (2013) by demonstrating the formation of associations between all elements of an engram. Furthermore, we extended the findings of Madan et al. (2010) that a single factor of an element, such as memorability, can strengthen the pairwise association of a memory.

In addition, the use of an episodic memory research paradigm offers greater ecological validity compared to other common research methods used in eyewitness testimony research. We instructed participants to vividly imagine that they were part of the crime event, thus forming episodic memories. This activates the hippocampus, which is involved in the formation of real-life memories (Horner & Burgess, 2013). As described by Erickson et al. (2022), the application of episodic memory research paradigms embodies great potential in the field of eyewitness memory research. This approach eliminates ethical risks and the need for expensive or specialised study environments.

Limitations and Further Directions

There are some limitations to the above-stated results. Firstly, there was an error in arranging the faces in one of the scenarios, which meant that the culprit's face was presented twice in one of the scenarios. Importantly, however, this error did not affect the comparison of accuracy for trials with memorable and non-memorable scenes because it occurred equally in both memorability conditions. Secondly, this study was conducted in an online environment. Therefore, we had no control over the participants' experimental conditions. This creates the potential for cheating, lack of concentration, or distractions such as loud noises. Even though, all participants performed above chance, suggesting engagement with the experiment, it could be beneficial to implement our research paradigm to the laboratory setting (Pause et al., 2013). In the laboratory setting one gains more control over environmental influences. In addition, other features of the crime event, such as characteristics of the eyewitness influence the reliability of eyewitness testimony. Thus, to assess the role of memorability on eyewitness reliability it would be necessary to evaluate possible interaction effects with other predictors of eyewitness reliability. For future research, other predictors of eyewitness reliability such as eyewitness intoxication, eyewitness emotionality or eyewitness confidence could be considered (Pezdek, 2012). Lastly, the practical implementation of our results is limited by the small effect size of the main effect of memorability.

In future research, we suggest further use of the episodic memory research paradigm to assess the reliability of eyewitness testimony. This would also allow for the analysis of information encoding and retrieval in episodic memory. We found evidence indicating associations between elements of an engram. Therefore, we argue in favor of the “Dependent Theory” by Horner and Burgess (2013). Nevertheless, we evaluated an alternative explanation for the dependency between associations in episodic memory retrieval. Further research is necessary to thoroughly assess the influence of sequential retrieval processes on eyewitness memory retrieval. An interesting consideration for further research is also the investigation of the cue type difference. It needs to be investigated whether scene images are inferior cue images due to scene-specific image properties. Additionally, one could use semantically related images of the object to see if the categorical clustering of information influences triplet encoding and retrieval.

Conclusion

This study shed light on the influence of the memorability of a crime scene on the reliability of eyewitness memory. Crime scenes with higher memorability can lead to greater accuracy in culprit identification. We implemented an episodic memory research paradigm into the field of eyewitness testimony research. This allowed the analysis of the “Dependent Theory“ of engram retrieval stated by Horner and Burgess (2013). The data suggests that the elements of an engram are retrieved interdependently. Further, we assessed the robustness of the dependency analysis, by exploring an alternative explanation for the data. This analysis was not able to falsify our initial results. Due to the extensive prior research (Horner & Burgess, 2013; Horner et al., 2014) and the exploratory analysis of our data, we conclude the dependency between associations of an engram. The elements seem to be associated with one another and the effect of the memorability of one element translates to the association between other elements of the same engram (Madan et al., 2010). The present study introduces a novel perspective by exploring the effects of memorability on eyewitness testimony. This research contributes to the expanding body of knowledge on image memorability and episodic encoding, while also shedding light on the reliability of eyewitness testimony.

References

Bainbridge, W. A., Isola, P., & Oliva, A. (2013). The intrinsic memorability of face photographs. *Journal of Experimental Psychology: General*, *142*(4), Article 4. https://doi.org/10.1037/a0033872

Borkin, M. A., Vo, A. A., Bylinskii, Z., Isola, P., Sunkavalli, S., Oliva, A., & Pfister, H. (2013). What Makes a Visualization Memorable? *IEEE Transactions on Visualization and Computer Graphics*, *19*(12), Article 12. https://doi.org/10.1109/TVCG.2013.234

Deffenbacher, K. A., Bornstein, B. H., Penrod, S. D., & McGorty, E. K. (2004). A Meta-Analytic Review of the Effects of High Stress on Eyewitness Memory. *Law and Human Behavior*, *28*(6), 687–706. https://doi-org.proxy-ub.rug.nl/10.1007/s10979-004-0565-x

Doyle, J.M. (2014). Eyewitness Evidence: Historical Background. In: Encyclopedia of Criminology and Criminal Justice, edited by Bruinsma, G., Weisburd, D., 277- 285. New York, NY: Springer. https://doi.org/10.1007/978-1-4614-5690-2_620

Dysart, J. E., Lindsay, R. C. L., MacDonald, T. K., & Wicke, C. (2002). The intoxicated witness: Effects of alcohol on identification accuracy from showups. *Journal of Applied Psychology*, *87*(1), 170–175. https://doi-org.proxy-ub.rug.nl/10.1037/0021-9010.87.1.170

Erickson, W. B., Wright, A., & Naveh-Benjamin, M. (2022). “He was the one with the gun!” Associative memory for white and black faces seen with weapons. *Cognitive Research: Principles and Implications*, *7*(1), 8. https://doi.org/10.1186/s41235-022-00355-z

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). GPower 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi-org.proxy-ub.rug.nl/10.3758/BF03193146

Ghasemi, A., Zahediasl, S. (2012). Normality tests for statistical analysis: a guide for non statisticians. Int J Endocrinol Metab. Spring;10(2):486-9. doi:10.5812/ijem.3505. Epub 2012 Apr 20. PMID: 23843808; PMCID: PMC3693611.

Horn, S. S., Bayen, U. J., & Michalkiewicz, M. (2021). The development of clustering in episodic memory: A cognitive-modeling approach. *Child Development*, *92*(1), 239–257. https://doi-org.proxy-ub.rug.nl/10.1111/cdev.13407

Horner, A. J., & Burgess, N. (2013). The associative structure of memory for multi-element events. *Journal of Experimental Psychology: General*, *142*(4), Article 4. https://doi.org/10.1037/a0033626

Horner, A. J., & Burgess, N. (2014). Pattern Completion in Multielement Event Engrams. *Current Biology*, *24*(9), 988–992. https://doi.org/10.1016/j.cub.2014.03.012

Horner, A. J., Bisby, J. A., Bush, D., Lin, W.-J., & Burgess, N. (2015). Evidence for holistic episodic recollection via hippocampal pattern completion. *Nature Communications*, *6*(1), 7462. https://doi.org/10.1038/ncomms8462

Isola, P., Xiao, J., Torralba, A., & Oliva, A. (2011). *What makes an image memorable?* 8.

Jones, G. V. (1976). A fragmentation hypothesis of memory: Cued recall of pictures and of sequential position. *Journal of Experimental Psychology: General*, *105*(3), 277–293. https://doi.org/10.1037/0096-3445.105.3.277

Khosla, A., Raju, A. S., Torralba, A., & Oliva, A. (2015). Understanding and Predicting Image Memorability at a Large Scale. *2015 IEEE International Conference on Computer Vision (ICCV)*, 2390–2398. https://doi.org/10.1109/ICCV.2015.275

List, J. A. (1986). Age and schematic differences in the reliability of eyewitness testimony. *Developmental Psychology*, *22*(1), 50–57. https://doi.org/10.1037/0012-1649.22.1.50

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory*, *4*(1), 19–31. https://doi-org.proxy-ub.rug.nl/10.1037/0278-7393.4.1.19

Madan, C. R., Glaholt, M. G., & Caplan, J. B. (2010). The influence of item properties on association-memory. *Journal of Memory and Language*, *63*(1), Article 1.

https://doi.org/10.1016/j.jml.2010.03.001

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law*, *7*(1), 3–35. https://doi-org.proxy-ub.rug.nl/10.1037/1076-8971.7.1.3

Meiser, T., & Bröder, A. (2002). Memory for multidimensional source information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(1), 116–137. https://doi.org/10.1037/0278-7393.28.1.116

Moult, M. (2011). Cued Recall. In: Kreutzer, J.S., DeLuca, J., Caplan, B. (eds) Encyclopedia of Clinical Neuropsychology. Springer, New York, NY. https://doi.org/10.1007/978-0-387-79948-3_1116

Pause, B. M., Zlomuzica, A., Kinugawa, K., Mariani, J., Pietrowsky, R., & Dere, E. (2013). Perspectives on episodic-like and episodic memory. *Frontiers in Behavioral Neuroscience*, *7*. https://doi-org.proxy-ub.rug.nl/10.3389/fnbeh.2013.00033

Pezdek, K. (2012). Fallible eyewitness memory and identification. In B. L. Cutler (Ed.), *Conviction of the innocent: Lessons from psychological research.* (pp. 105–124). American Psychological Association. https://doi-org.proxy-ub.rug.nl/10.1037/13085-005

Pilotti, M., Beyer, T,. Yasunami, M. (2001). Encoding Tasks and the Processing of Perceptual Information in Young and Older Adults. *The Journals of Gerontology: Series B*, 56, 119 128. https://doi.org/10.1093/geronb/56.2.P119

Robinson, M. D., Johnson, J. T., & Herndon, F. (1997). Reaction time and assessments of cognitive effort as predictors of eyewitness memory accuracy and confidence. *Journal of Applied Psychology*, *82*(3), 416–425. https://doi-org.proxy-ub.rug.nl/10.1037/0021-9010.82.3.416

Sporer, S. L., Penrod, S., Read, D., & Cutler, B. (1995). Choosing, confidence, and accuracy: A meta-analysis of the confidence-accuracy relation in eyewitness identification studies. *Psychological Bulletin*, *118*(3), 315–327. https://doi-org.proxy-ub.rug.nl/10.1037/0033-2909.118.3.315

Tulving, E. (1984). Précis of Elements of episodic memory. *Behavioral and Brain Sciences*, *7*(2), 223–268. https://doi-org.proxy-ub.rug.nl/10.1017/S0140525X0004440X

Tulving, E. (1985). How many memory systems are there? *American Psychologist*, *40*, 385–398. https://doi.org/10.1037/0003-066X.40.4.385

Wulff, A. N., & Hyman Jr, I. E. (2022). Crime blindness: The impact of inattentional blindness on eyewitness awareness, memory, and identification. *Applied Cognitive Psychology*, *36*(1), 166–178. https://doi.org/10.1002/acp.3906

Xie, W., Bainbridge, W. A., Inati, S. K., Baker, C. I., & Zaghloul, K. A. (2020). Memorability of Words in Arbitrary Verbal Associations Modulates Memory Retrieval in the Anterior Temporal Lobe. *Nature Human Behaviour*, *4*(9), 937–948. https://doi.org/10.1038/s41562020-0901-2

Apendendix A

Table and Figure of The Method Section

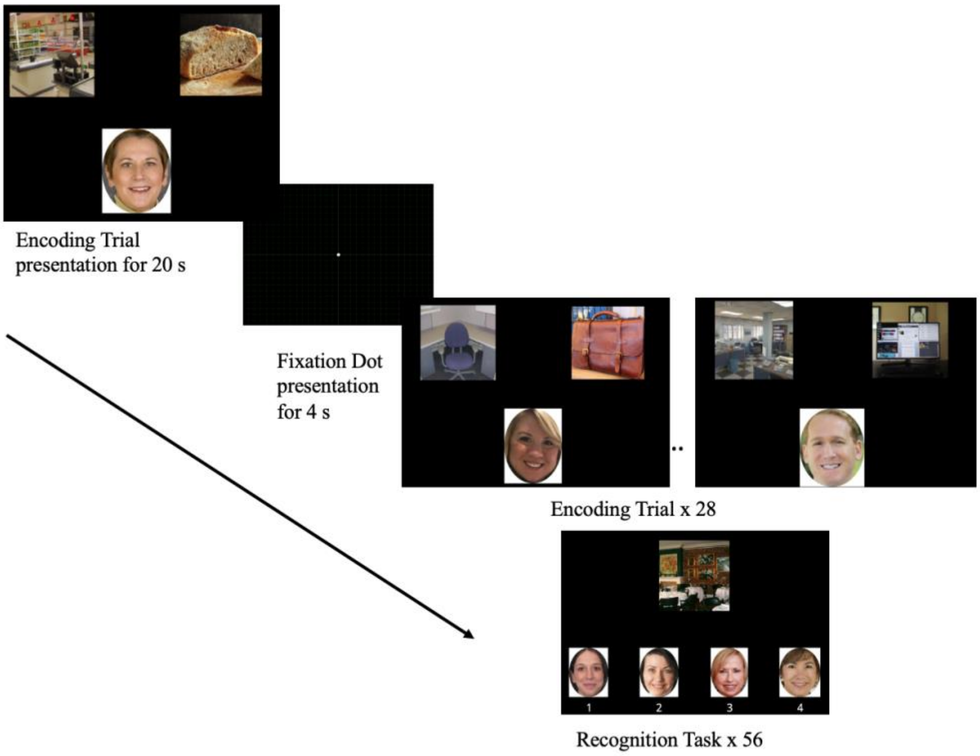
Table A1

Scene Types and Associated Objects

Scene Description	Assigned Objects
Attic	CD Player/Luggage
Bakery	Sandwich/Cash Register
Barrel Basement	Flask/Ashtray
Bedroom	Book/Piggy Bank
Canteen	Plate/Shopping Bag
Church	Candelabra/Cross
Hallway	Doormat/Lightbulb
Kitchen	Mixer/Toaster
Living Room	Television/Telephone
Music Room	Amplifier/Trumpet
Storage Room	Box/Pallet
Supermarket	Money/Bread
Weapon Museum	Backpack/Pocket Watch
Workspace	Briefcase/Computer

Figure A1

Visual Representation of the Experimental Procedure

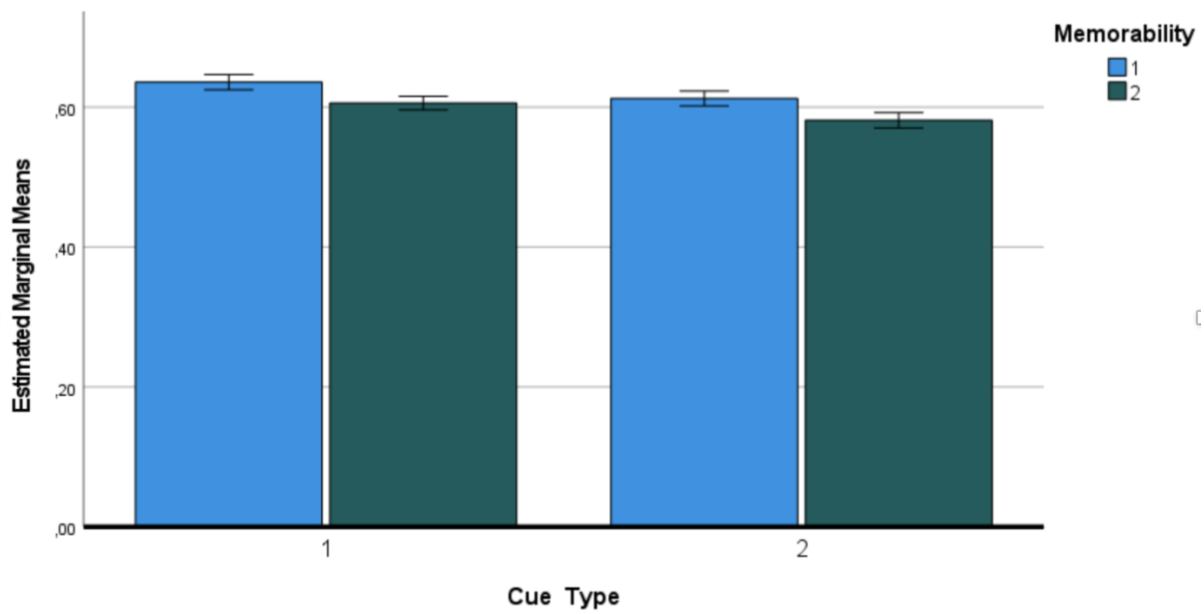


Appendix B

Data Visualisation

Figure B1

Memorability Effect on Culprit Identification Accuracy Split by Cue Type

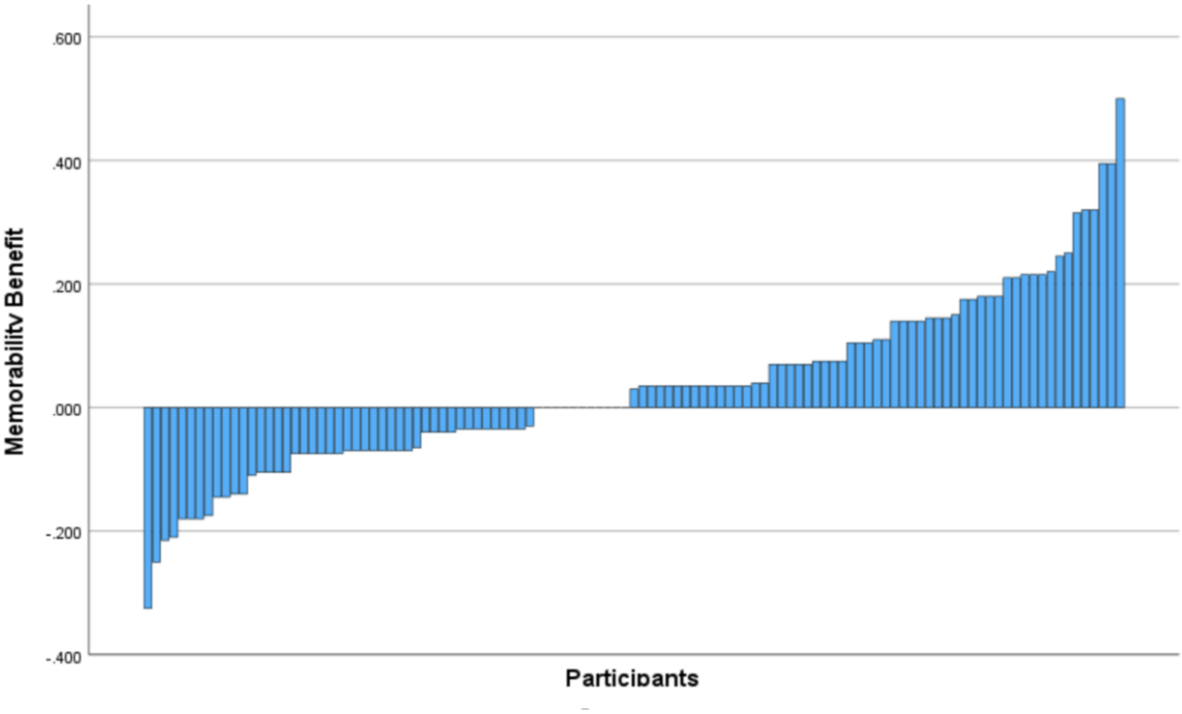


Note: The dependent variable, identification accuracy is displayed on the Y-axis. On the x-axis is the cue type displayed. Cue Type one is the object condition and two is the scene condition. The different bars distinguish the two memorability conditions (memorable (one) vs non-memorable (two)). The graph shows the difference in identification accuracy under memorable and non-memorable scene conditions. Further, this difference is visible in both cueing conditions (object and scene).

**Note: The error bars display one standard error of the within-subject error. We normalized the data to control for between-subject errors.*

Figure B2

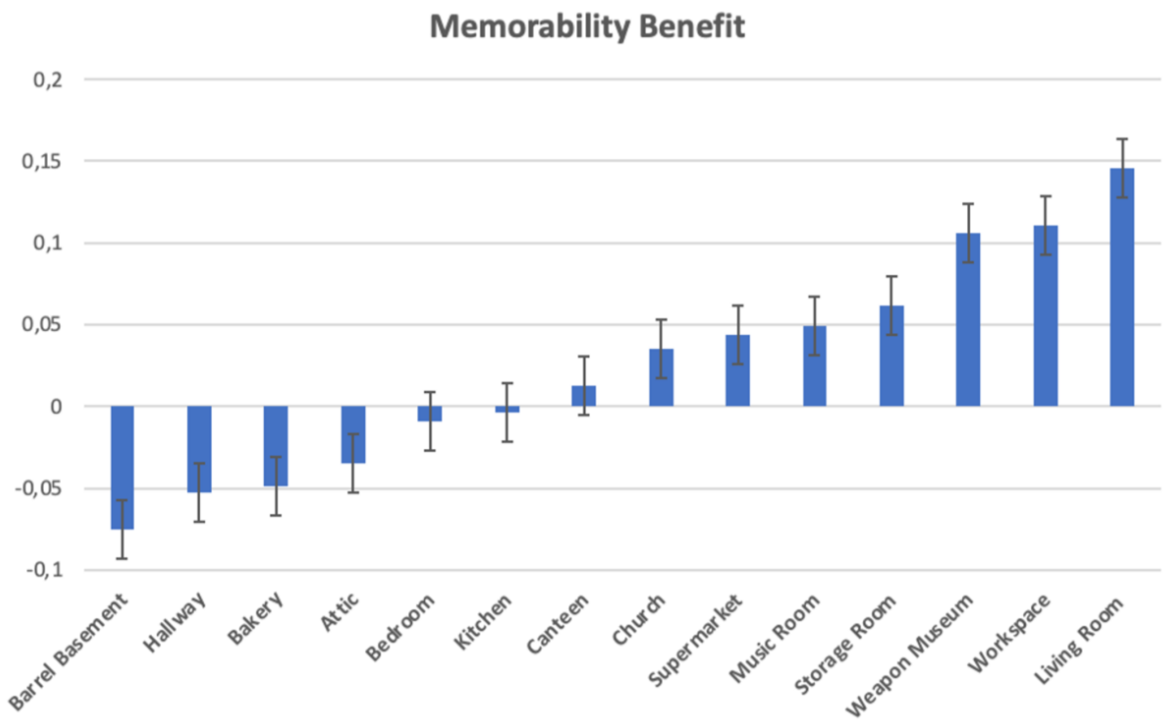
Average Memorability Benefit for each Participant



Note: This graph displays the memorability benefit score for each participant. The participants are ordered by the memorability benefit score from lowest to highest. The majority of the sample displays a positive memorability benefit score, indicating higher culprit identification in the memorable scene condition than in the non-memorable scene condition.

Figure B3

Average Memorability Benefit per Scene

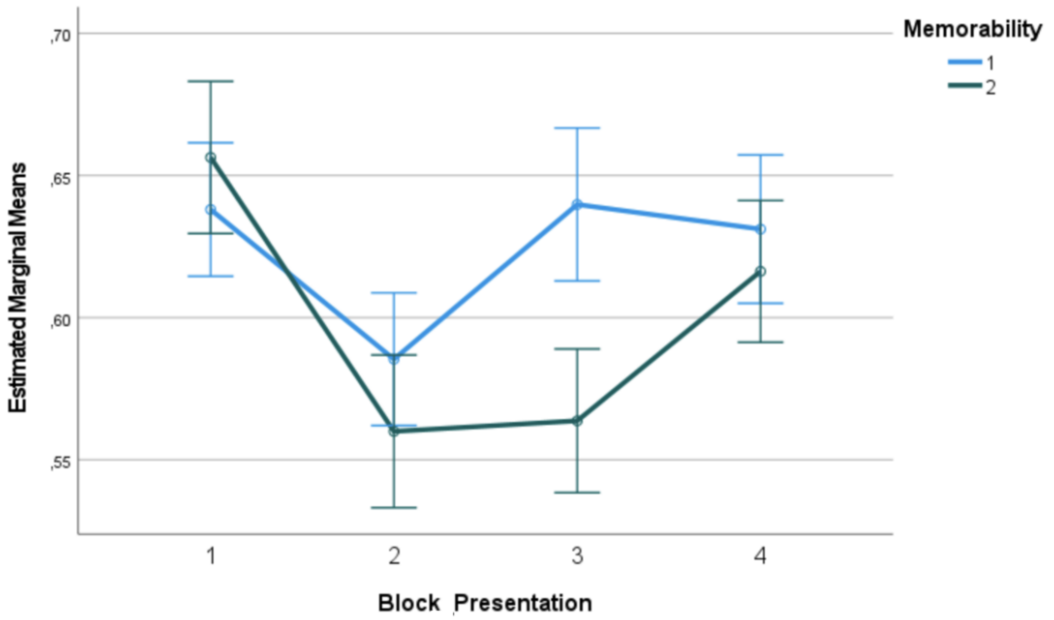


Note: This graph displays the memorability benefit score for each crime scene. The scenes are ordered by the memorability benefit score from lowest to highest. The majority of the scenes display a positive memorability benefit score, indicating higher culprit identification in the memorable scene condition than in the non-memorable scene condition.

Note: The error bars display on standard error.*

Figure B4

Memorability Effect on Culprit Identification Accuracy Across Presentation Blocks

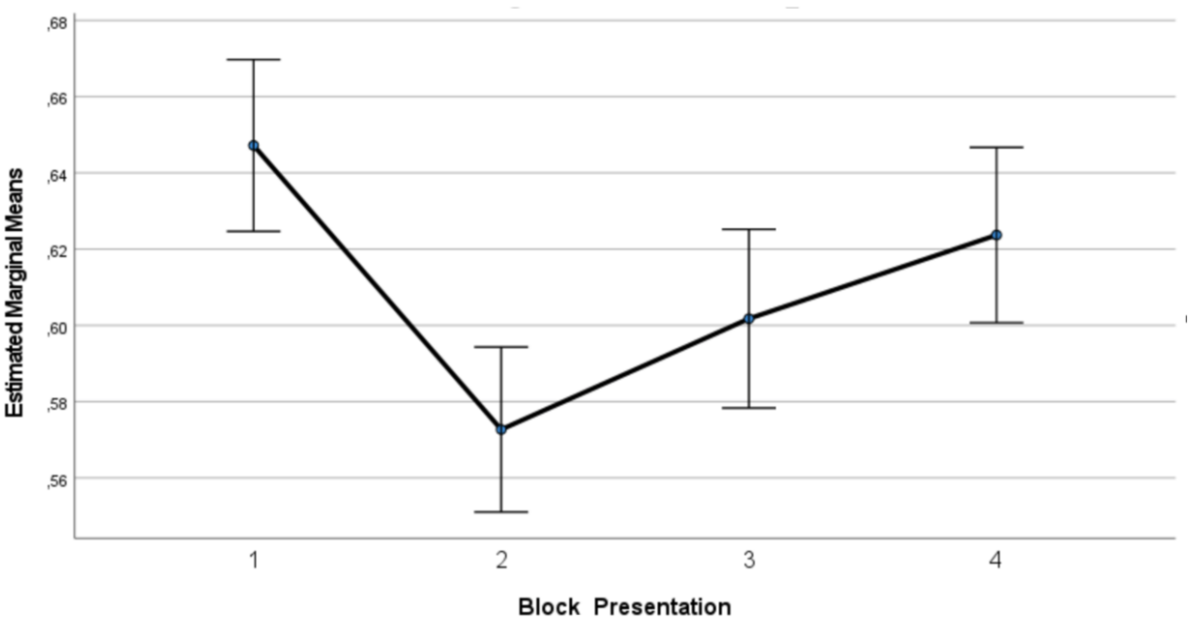


Note: The dependent variable, identification accuracy is displayed on the Y-axis. On the x-axis is the presentation block displayed. The different lines distinguish the two memorability conditions (memorable (one) vs non-memorable (two)). The graph shows the difference in identification accuracy under memorable and non-memorable scene conditions across presentation blocks.

Note: The error bars display on standard error.*

Figure B5

Average Identification Accuracy Across Presentation Blocks



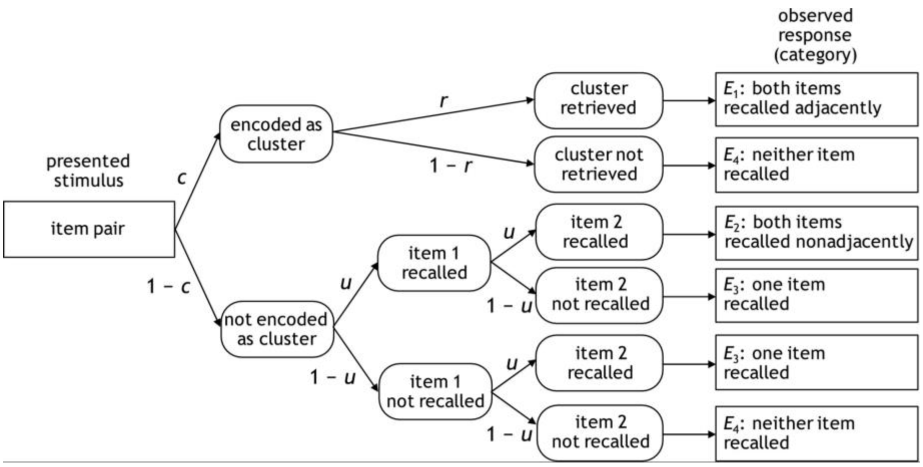
Note: The dependent variable, identification accuracy is displayed on the Y-axis. On the x-axis is the presentation block displayed. The graph shows the average identification accuracy across presentation blocks.

Note: The error bars display on standard error.*

Appendix C

Model Visualisation

The Pair-Clustering Multinominal Model



Note: *Reprinted from "The Development of Clustering in Episodic Memory: A Cognitive-Modeling Approach". Horn, S. S., Bayen, U. J., & Michalkiewicz, M. (2021). The development of clustering in episodic memory: A cognitive-modeling approach.* Retrieved from https://doi-org.proxy-ub.rug.nl/10.1111/cdev.13407