Carreras, F., Moulin, C. J. A., Tales, A., Barnes, C. M., & Souchay, C. (2024). Metacognitive processes accompanying the first stages of autobiographical retrieval in the self-memory system. Memory, 1–14. http://dx.doi.org/10.1080/09658211.2024.2370532

Metacognitive Processes Accompanying the First Stages of Autobiographical Retrieval in the Self-Memory System

Fabien Carreras[1, 2, 3]; Chris J. A. Moulin[1]; Andrea Tales[2]; Claire M. Barnes[3]; Céline Souchay[1]

[1]Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LPNC, 38000 Grenoble, France
[2]Centre for Ageing and Dementia Research (CADR), Swansea University, Swansea, SA2 8PP, Wales, UK
[3]Department of Biomedical Engineering, Swansea University, Swansea SA1 8EN, Wales, UK

Corresponding author fabien.carreras@univ-grenoble-alpes.fr
Address: 1251 rue des Universités CS 40700, 38058 Grenoble Cedex 9 France

Source(s) of support: This work was supported by a studentship funded by the ANR-15-IDEX-02 METASTORY project awarded to Céline Souchay (Université Grenoble Alpes), Claire Barnes and Andrea Tales (Swansea University).

Word count: 10874
Word count including abstract: 11070

ORCIDs:
Fabien 0000-0002-3073-0341
Chris 0000-0001-9784-4362
Céline 0000-0001-9041-0123
Andrea 0000-0003-4825-4555
Claire 0000-0003-1031-7127

Abstract

According to Conway's view, Autobiographical memory (AM) construction is accompanied by control processes. These processes range from filtering out relevant memories according to the current context, to generating or elaborating appropriate retrieval cues. These processes can be conceptualised as metacognition, the ability to control and monitor cognitive processes. Experimentally, little has been carried out to support the idea that metacognition is involved in AM. To assess this, we designed a task, the Feeling of Retrieval. Participants had to predict whether cue words would facilitate AM access (i.e., fluent access cues) or not (i.e., limited access cues) in a limited time (either 1 (Exp. 2) or 2 (Exp. 1) s). Later, they retrieved memories in response to both types of cues. Results show that cues judged as fluent access led to better AM generation, as illustrated by AM retrieval latency and a subjective measure of the ease with which the AMs were retrieved. These rapid predictions may rely on epistemic feelings and / or other mnemonic cues such as a partial retrieval of information. This metacognitive access to the earliest stages of AM retrieval illustrates the ability to monitor AM processes as proposed by Conway (2005).

Keywords
Autobiographical Memory, Retrieval Mode, Metacognition, Feeling of Fluency, Feeling of Retrieval.

Introduction

How do we know that we have an autobiographical memory (AM) in mind? Although theoretical works stress the involvement of metacognition in the retrieval of autobiographical memories (e.g., Mazzoni & Kirsch, 2002; Moulin et al., 2022; Welch-Ross, 1995) – or more broadly the presence of control processes and structures that shape retrieval (e.g., the working self, Conway et al., 2019) – and the implication of AM in metacognitive processes (e.g., Morris & Mograbi, 2013), experimental investigations of metacognitive processes related to AM have been carried out only recently (e.g., Carreras & Moulin, 2023; Matsumoto, 2022). In this article, we aim to test recent assumptions that metacognition guides AM retrieval (Barzykoswki & Moulin, 2022). Investigating how metacognition influences AM processes at the earliest stage would allow us to better understand AM retrieval. Here, we briefly review the recent literature about AM retrieval within the Self-Memory System before giving an overview of metacognition and describing its role in AM retrieval.

According to the Self-Memory System (SMS; Conway & Pleydell-Pearce, 2000; Conway, 2005; Conway et al., 2019), AM is organised in a hierarchical structure. The top levels of this structure are semantic representations of one's life or periods of life. The lower the components are in the hierarchy, the more they are episodic, and eventually are episodic representations of events one's experienced, the specific autobiographical memories. Retrieval occurs either in a bottom-up or top-down fashion in this hierarchy, either driven by spontaneous retrieval of episodic detail, or as a strategic research process undertaken when entering the retrieval mode and guided by the knowledge and self-structures.

Conway (2005) proposed that such voluntary retrieval can be either direct or generative. In direct retrieval, a cue activates a specific memory, and this activation is spread to higher levels of the AM hierarchy. In the generative retrieval, a cue activates a high level of the AM structure, this activation is spread to more specific AM components underlying a voluntary research in memory aligning with the SMS retrieval mode in which individuals create expectations about their retrieval (Conway et al., 1999; Conway, 2005). Research has shown that people can reflect on their retrieval process while doing this search within the hierarchy (Uzer et al., 2012). Recently, Mace et al. (2021) added three other types of AM generative retrieval: temporal recall strategies, repeating strategies and visual imagery strategies. Importantly, these findings are inherently metacognitive: participants are able to reflect on their AM retrieval processes, that is, they describe and express their retrieval while doing it. This shows that by definition, generative retrieval is metacognitive and overlaps with awareness since individuals can, at some point, monitor their retrieval, be conscious of and report how they are accessing their memories.

A recent view of AM aimed at elucidating retrieval mechanisms within the SMS by proposing a threshold account (Barzykowski et al., 2019; Barzykowski & Staugaard, 2016, 2018). This threshold hypothesis suggests that each specific AM has a threshold. Activation is constantly spread across memories, and when a memory has enough activation, it may enter consciousness. The accessibility of a memory depends on its characteristics. AMs that are the most vivid, emotionally intense, important, rehearsed and unique are more easily accessed; all these are consistent with having a higher level of activation and therefore mean they are more likely to pass the threshold. In this view, the threshold of a specific AM can be lowered when one enters what Conway (2005) referred to as the retrieval mode, in which they are expecting a memory retrieval and start an active search. The intention to retrieve an AM and the monitoring of thoughts are two metacognitive processes that influence the AM threshold.

According to the threshold hypothesis, AM retrieval has four stages (Barzykowski & Mazzoni, 2021; Barzykowski & Moulin, 2022; Moulin et al., 2022). The first is the pre-retrieval stage. In this stage, individuals create expectations about the sort of memory they will retrieve (e.g., a memory from their high school), and enter the retrieval mode (Conway & Pleydell-Pearce, 2000). Priming effects occur here: Cues can lower the memory threshold of memories that are linked to them. That is, some memories will have higher chances to enter awareness. Also, people may monitor their thoughts to focus on AM. Monitoring and priming effects occurring at this stage may facilitate or impair retrieval according to the threshold modifications provoked. The second stage is the retrieval stage. In this stage, AMs are triggered and / or accessed via a given cue. Here, as postulated by the SMS, AMs can be (re)constructed without awareness or directly accessed. In the post-retrieval stage, the memory formed reaches consciousness. If the retrieval is voluntary, and the memory meets the given criteria, the retrieval can be terminated. The final stage is the retrieval outcome report stage. Here, the memory can be verbally shared or not.

Both direct and generative retrieval are most often triggered by cues. The feeling of familiarity accompanying cue processing has been discussed as a variable influencing AM retrieval. Recently, Moulin et al. (2022, see also Barzykowski & Moulin, 2022) have suggested that the feeling of familiarity triggered by a cue can provoke or inhibit a retrieval attempt. They view familiarity as metacognitive: familiarity would be the result of an interpretation of the fluency with which a cue is processed, something which is known to influence retrieval in episodic memory more generally (e.g., Bastin et al., 2019).

In this framework, familiarity is a metacognitive feeling that guides AM retrieval. Proust (2007; and later Arango-Muñoz, 2011) suggested that metacognition entails two levels. The higher level comprises metarepresentations (as suggested by Nelson and Narens, 1990) which gives us the ability to make judgements on our own and other's mental states. Metacognitive judgements made at this level are interpretations of and inferences about cognitive processes. The lower level of metacognition is driven by epistemic feelings; subjective and spontaneous experiences occurring in response to cues or situations (e.g., a feeling of familiarity when seeing an old friend after years; Arango-Muñoz, 2014). Epistemic feelings are informative but do not involve any metarepresentation, they inform us about how our cognition is doing at the time and allow us to monitor it (Michaelian & Arango-Muñoz, 2014).

Metacognitive monitoring has often been investigated in memory tasks, but rarely in AM. It has been tested using two methods: the retrospective evaluation of the performance, and the prediction of upcoming performance (Fleming and Lau, 2014). We took as inspiration for the current article the feeling of knowing (FOK) paradigm in which participants predict their upcoming performance. A common FOK procedure uses cue-target pairs (e.g., Koriat & Levy-Sadot, 2001). Participants first learn pairs of words. Then, they carry out a recall task in which they see a cue and have to recall the second word of the pair. When they fail to recall an item, they have to predict whether they will recognise it or not; the FOK judgement. Then there is a final memory test. Participants make accurate FOK judgements (e.g., Morson et al., 2015).

The interactive hypothesis was put forward to explain the nature of FOK judgements (Castel et Middlebrooks, 2016; Koriat & Levy-Sadot, 2001). This states that a two-stage procedure intervenes in the FOK. First, cues create a feeling of familiarity. The more this epistemic feeling is strong, the more it leads to a sensation of target retrievability and eventually to a retrieval attempt (Metcalfe et al., 1993; Reder, 1987). Second, information accessed through the retrieval attempt would be the basis for the FOK judgement (Koriat, 1993). These two stages correspond to the two levels of metacognition proposed by Arango-Muñoz (2011). First, behaviours are monitored by epistemic feelings (i.e., here a feeling of familiarity may lead to a retrieval attempt), then, it is monitored by an inference based on an evaluation of current mental states. In line with the interactive hypothesis and as suggested by Moulin et al. (2022, see also Barzykowski & Moulin, 2022), we propose that AM voluntary retrieval is triggered by epistemic feelings such as familiarity. Here we aim to extend the use of the FOK procedure to test this assumption.

Recently, two studies have used a combination of AM and metacognitive procedures to assess metacognitive abilities related to AM (Carreras & Moulin, 2023; Matsumoto, 2022). Using a prediction performance procedure, Matsumoto (2022; in pre-print) asked his participants to indicate whether each cue from a sample could facilitate the retrieval of a specific AM on a 1-7 Likert Scale. Also, participants had to indicate the valence of the memory the cue word would evoke. Next, participants retrieved memories, specific or not, in response to the cue words. Results showed a positive link between the ease of retrieval judgement and the type of AM retrieval. The easier the retrieval was predicted in response to words, the more often the retrieval was direct (i.e., fast and effortless). Furthermore, the more the anticipated emotion was positive, the more the retrieval was direct.

In this current article, we designed a new task asking participants to predict their performance in an AM task, the Feeling of Retrieval (FOR). This task is similar to Matsumoto's (2022) ease of retrieval task, but was inspired by Reder and Ritter's (1992) FOK procedure. In their procedure, they presented an arithmetic problem to participants, who had 850 ms to indicate if it would be easy or not to solve it. Depending on their answer, participants respectively had 1 sec to solve it or 18 sec to work on it and suggest a solution. This procedure showed that participants were able to quickly predict their performance, and it emphasised the role of familiarity in FOK judgements. We adapted this procedure to autobiographical memory. Of note, we are not the first to explore AM with reaction time measures. Conway and Bekerian (1987) looked at the different reaction times taken to respond to material from different levels of the SMS hierarchy in an approach which was inspired by Rosch's approach to investigating semantic memory (1975).

In our FOR, participants see cue words on the screen. Each word is presented for a limited time, one at a time. For each, participants indicate if the word would be an easy cue for which to retrieve a specific AM. In other words, they judge how easy it would be to perform a Galton-Crovitz cueing task for each word they see. Judgements are made in a speeded go-no go manner, rather than with an explicit judgement for each word, i.e. participants press a button as quickly as possible if they think the cue could facilitate AM retrieval. This limited time should force participants to respond on the basis of metacognitive feelings rather than on a complete AM retrieval. After these predictions, in a later task, they perform a standard AM retrieval task with a sample of words randomly selected by the experimenter according to whether or not the word was identified as allowing fluent access to their AM (i.e., *fluent access cues*) or not (i.e., *limited access cues*).

We thus designed the FOR task to assess monitoring at the pre-retrieval stage. That is, the ability to judge on-line the effectiveness of a cue-word in priming participants' AM. Barzykoswki and Moulin (2022) suggested that epistemic feelings occurring in the early stages of retrieval guide AM retrieval. If such metacognitive processes exist in AM retrieval, they should be used in the limited time participants have to make their predictions. Accordingly, participants would base their prediction on similar feelings as observed in the first stage of FOK predictions. If participants have metacognitive access to this very early stage of autobiographical retrieval, they should retrieve memories more fluently for cues they previously categorised as fluent access cues. Participants should retrieve memories faster and rate the retrieval as easier in response to fluent access cues. With retrieval being easier, fluent access cues should be associated with memories that have higher levels of activation. In line with the threshold hypothesis, memories retrieved in response to fluent access cues should therefore be more vivid, personally significant, specific, and emotionally intense.

Our other hypotheses are related to the AM task, where we took subjective ratings of typical characteristics in the empirical literature on AM. First, we aim to replicate a common finding that the faster the memories are retrieved, the more the memories are vivid, personally significant, specific and positive (Barzykowski and Mazzoni, 2022; Barzykowski et al., 2019). Second, we expect that the more fluently retrieved the memory is in terms of an objective measure of RT, the higher the subjective rating of fluency will be. Third, we assess whether retrieval fluency is linked to memory ratings, and correlations between the different memory ratings. We also explored whether the speed with which participants categorised words as fluent access cues in the prediction task was linked to AM retrieval fluency and memory ratings. Finally, we explored whether the properties of the words used in the experiments could influence participants' classification and performance.

We carried out two studies using this paradigm. Healthy adults were recruited for both studies. In the first, participants had 2 s to make their prediction. In the second, they only had 1 s to do so. These studies were pre-registered in the Open Science Framework. All data and scripts are available online (https://osf.io/r2j74/?view_only=d19767270c294be1b50b0c3bb53db43c).

Experiment 1

Method

Participants and Materials. The study was approved by the local institutional ethics committee (CERGA-Avis-2023-19). We ran a power analysis based on data from a pilot study (n = 19) and the simulation of data from 100 participants (*simr* R package version 1.0.6; Green and MacLeod, 2016). Our power analysis was carried out using the linear mixed model designed to test our main hypothesis, that is, participants predict correctly which words will facilitate autobiographical memory generation. This power analysis revealed that a sample size of 40 participants would allow us to reach a power of .80. Since our design is new and no data is available to estimate an effect size from a bigger sample, we aimed to reach a power of .90. That is, we aimed to analyse data from 57 participants.

We recruited 62 participants (M_{age} = 20.67, SD_{age} = 5.93, 73% of female). Two of them failed the attention checks (described later), leading to a sample size of 60 participants (M_{age} = 20.69, SD_{age} = 6.03; 82% of female). This study was divided into four phases: a training phase followed by three memory tasks (Figure 1).

This study was implemented on Psychopy (version 2022.2.4; Peirce et al., 2019). Data collection was carried out in person. Participants were tested individually in a small room. Before starting the experiment, all participants gave their consent to participate. All the words used in this experiment were randomly selected from Miceli et al. (2021). This database contains words associated with several properties, notably, word familiarity, imageability, concreteness and age of acquisition.

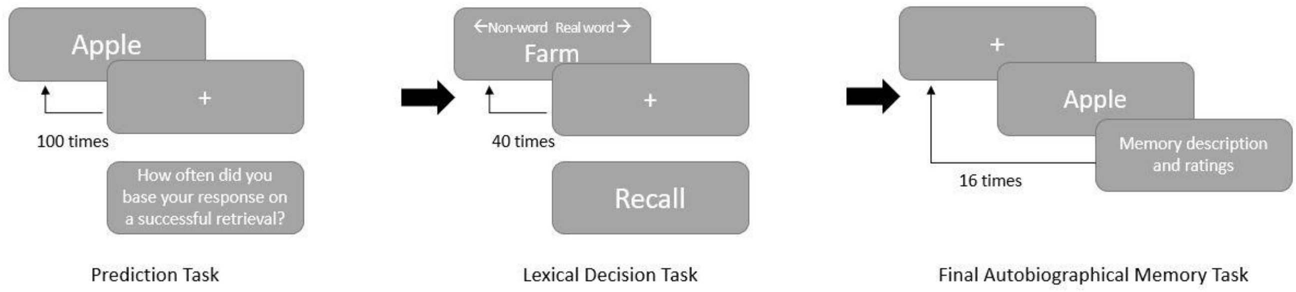

Figure 1. Illustration of the three memory tasks.

Procedure.

Training phase. Training was a typical cue-word AM task. Participants saw the cue word ball and were instructed to retrieve a specific AM, related or not to the word. They were told that it may or may not be a memory related to an event they think of as important, but it should have happened at a specific point in their life, whether recent, long ago, or at any other time. The memory had to be related to a specific event with a duration of less than 24-hours. These characteristics were explained to participants and examples of AMs were given.

While the cue word was on the screen, participants had to press the right arrow of the keyboard as soon as they thought of a memory. Next, they had to describe their memory to the experimenter in a few sentences and to rate it on several dimensions: vividness, importance, and specificity (if the memory was related to a specific place, location, and event). Also, they were asked to indicate how easy it was to generate the memory. Every rating was carried out on a 1-7 scale, with 7 being the highest rating (i.e., 1 corresponds respectively to not vivid at all, not important at all, not specific at all, not easy at all; and 7 corresponds respectively to extremely vivid, extremely important, extremely specific, extremely easy). Also, participants rated the valence of the memory on a (-3) - (+3) scale, with (-3) being extremely negative and (+3) being extremely positive.

This training phase had two goals. First, it aimed to facilitate the comprehension of the second task of the study. Second, participants were asked to type out their memory description and then present it. This allowed the experimenter to tell the participant if the description corresponds to an AM or not. If not, the experimenter was able to give feedback to ensure that the participant understood the type of memory to retrieve (e.g., this is not a specific AM; this is not related to a unique event…)

Prediction Task. In this task, participants were exposed to word stimuli on the screen. The stimuli were 100 words (again French nouns randomly selected from Miceli et al. (2021)), and 6 attention checks. The same 100 words were used with all participants. Attention checks were the sentence “do not press” and the sentence “press the Space bar”. Each sentence appeared 3 times on the screen at random intervals.

Each stimulus appeared on the screen, one at a time. Each time a cue word appeared on the screen, participants had to choose either to press the space bar on the keyboard, or not to respond. They had to press the space bar as quickly as possible if they thought that the cue on the screen would facilitate the generation of an AM, in other words, they judged if the cue word would easily lead to the recall of a specific AM. They did not respond if they thought that the word present on the screen would not lead to them readily generating an AM (i.e., a go-no go methodology). Thus, every participant categorised the cue words into two types: the “fluent access cues”, cue words associated with a response, which were supposed to lead to AM generation, and the “limited access cues", words for which participants did not respond, which were less likely to facilitate AM generation. Importantly, each participant had their own list of “fluent access cues” and “limited access cues” words based on their responses in the prediction task. They were told to not try to retrieve AMs to make their decisions.

In this first experiment, each stimulus remained on the screen for 2 seconds, regardless of the participant's response. This presentation time was based on feedback and data collected with our pilot study. Presentation of the cue word was followed by a fixation cross which remained on the screen between 0.5 and 1 s. The presentation time was random in order to keep participants concentrated on the task. Participants had to respond while the word was on the screen. Nothing happened if they pressed the button while the cross was on the screen. Participants had the opportunity to take a break after seeing half of the stimuli.

At the end of the prediction task, participants were asked whether they had voluntarily tried to generate AMs. Participants who answered yes were asked how often they based their answers on whether or not they were able to retrieve a memory. They responded on a 1-100 scale, with 100 being “100% of the time”. These questions examined whether participants used a retrieval strategy to answer the question, and no assumptions were made about their influence on the results.

Lexical Decision Task. The lexical decision task served as a distractor task, designed in order to make participants somewhat forget about their responses in the prediction task. Participants saw 40 strings of letters. For each, they had to indicate if it was a real word (20 trials), or a non-word (20 trials). Also, they were asked to remember the real words in order to recall these at the end of this task. Real words used in this task were selected from Miceli et al. (2021). These words were not presented during the prediction task nor for the training part. These words were the same for each participant. Participants responded using the left and right arrows of the keyboard. They had to press the left arrow to categorise the string of letters as a non-word, and the right arrow to categorise it as a real-word. They were asked to answer as quickly and as precisely as possible. After categorising the 40 stimuli, they had to recall as many real-words as possible.

Final Autobiographical Memory Task. The procedure was the same as in the training, except that it was carried out 16 times. Participants saw a word on the screen, either a fluent access cue or a limited access cue. They had to press the right arrow button as soon as they thought of a memory. Their response time was used as an indicator of the time they needed to access the memory (i.e., retrieval latency). They were instructed that it may or may not be a memory related to an event they think of as important, but it should have happened at a specific point in their life, whether recent, long ago, or at any other time. A memory had to be related to an event taking place within a 24-hour period. We asked them not to carry over the same memory several times. They retrieved memories in response to eight words they categorised as fluent access cues, and in response to eight words they categorised as limited access cues. These words were randomly sampled from each participant's list of fluent access cues and limited access cues. There was no time limit to generate memories. Participants were told to try for at least 1 minute to generate something. If nothing came up, they were instructed to press the button and write that they could not generate a memory.

After pressing the button, they were asked to write a description of their memory in a few sentences with no time limit. Then, they had to indicate how easy it was to access the memory on a 1-7 scale (with 1 being *Not easy at all* and 7 being *Extremely easy*) and to indicate the date of occurrence of the event. Also, they had to rate it on several dimensions: vividness, specificity, and importance. Each rating was carried out on a 1-7 scale, with 7 being the greatest rating (i.e., 1 corresponds respectively to not vivid at all, not important at all, not specific at all; and 7 corresponds respectively to extremely vivid, extremely important, extremely specific.

Also, participants had to rate the valence of the memory on a (-3) - (+3) scale, with (-3) being extremely negative and (+3) being extremely positive. We calculated emotional intensity by using the absolute value of the valence rating.

Data Preprocessing and Analyses. All thoughts generated during the AM task were read by the first author and a neutral judge. They categorised each description as a specific AM or as something else. They categorised the 30 first descriptions together and then they carried out this task alone and compared their responses. Disagreements were discussed and resolved so that all thoughts could be categorised. Agreement rate was 96% (Cohen's kappa = .83). Measures (e.g., generation time, vividness rating) that were not related to descriptions categorised as specific AM were removed from the analyses. Out of the 944 recorded descriptions, 826 have been categorised as memories (87.5%) and 188 have been categorised as something else (12.5%).

Participants who failed 3 or more attention checks were removed from the analyses (*n* = 2). Participants who did not retrieve AMs in response to words they categorised as limited access cues were removed from the analyses involving the type of cue (*n* = 4). For analyses involving response times, we removed all observations above or below 2.5 SD of the mean of the individual (*n* = 23).

Analyses were conducted using R (version 4.1.1; R Core Team, 2017). Some analyses involved linear mixed models. Each time we used linear mixed models, random effects relevance was tested using Bates et al.'s (2015) procedure and p-values were calculated using Satterthwhaite approximation, available in the *lmertest* R package (version 3.1-3; Kuznetsova et al., 2017). In parallel to the analyses presented below, we also analysed the data using RT as a measure using a log transformation and found that the significance of the results was unchanged.

Results

The results concern the relationship between the cues selected in the first phase and the ease and speed at which specific memories come to mind in the last phase. As such, we first investigate the differences between cues selected as fluent access or limited access, this being our main hypothesis. Second, we examine standard characteristics of the autobiographical memories, and in a final section we analyse the cue words to see which types of cues were most likely to generate AMs.

Effects of cue type on retrieval processes. We designed a linear mixed model with participants as a random factor to investigate the effect of the type of cue (i.e., limited access cue or fluent access cue) on AM retrieval latency. The type of cue was used as a predictor and the AM generation time as a measure: $b = -2.8305$, $t(721.6597) = -7.182$, $p < .001$. Overall, AMs were retrieved 2.8 s faster in response to fluent access cues than in response to limited access cues (Table 1).

We designed a linear mixed model with participants as a random factor to investigate the effect of the type of cue (i.e., limited access cue or fluent access cue) on the ease of generation rating. The type of cue was used as a predictor and the ease of generation rating as a measure: $b = 0.9414$, $t(55.3843) = 6.679$, $p < .001$. That is, participants rated the generation of an AM in response to a fluent access cue as easier than in response to a limited access cue.

To interpret this result, we need to explore the basis of this appropriate metacognitive classification. To do this we can examine whether or not participants used a retrieval strategy, i.e. started to access (or actually did access) a memory in the prediction task. In total, 49 (i.e., nearly all) participants responded that they had tried to retrieve AMs during the prediction task. On average, they indicated that they relied on an AM retrieval attempt 60% of the time. Using a linear model[1] with the type of cue and the frequency with which participants relied on their voluntary retrieval to categorise words as fluent access cues or limited access cues as predictors, we found that the effect of the type of cue on retrieval fluency was not modulated by the frequency with which participants used this strategy, $b = -0.006121$, $t(47) = -0.234$, $p = .816$, $PRE = 0.001$. We carried out the same analysis adding back in those participants who did not try to retrieve AMs and still found no difference: $b = - 0.01537$, $t(54) = -0.818$, $p = .417$, $PRE = 0.012$. Similarly, such a retrieval attempt did not modulate the effect of the type of cue on ease of generation ratings, both when the analysis contained only participants who tried to

[1] We calculated a mean of RT for fluent access cues, and a mean of RT for limited access cues for each participant and created a score which reflected the difference in RT between fluent access cues and limited access cues for each participant: this is the variable we used to assess type of cue. The linear model contained this score (Wdiff) as its measure, and the reliance on successful retrieval (β1) as its predictor: Wdiff~β1 (or $Wdiff = \beta0 + \beta1 + \epsilon$).

generate AMs, $b = 0.006601$, $t(47) = 0.992$, $p = .326$, $PRE = 0.02$ and when the analysis contained all participants, $b = 0.006428$, $t(54) = 1.355$, $p = .181$, $PRE = 0.033$.

We designed linear mixed models with the type of cue as predictor and with the memory ratings as measures to assess whether fluent access cues would lead to different AMs than limited access cues. Participants were the random factor of the models. We found that AMs retrieved in response to fluent access cues were rated as more vivid, $b = 0.65795$, $t(53.43618) = 5.273$, $p < .001$, more specific, $b = 0.5193$, $t(742.0530) = 4.841$, $p < .001$, more important, $b = 0.6510$, $t(742.7483) = 5.512$, $p < .001$ and more emotionally intense, $b = 0.25837$, $t(744.22770) = 3.594$, $p < .001$. However, we found no effect of the type of cue on AM valence, $b = 0.1161$, $t(748.1546) = 1.007$, $p = .314$.

Finally, we designed a linear mixed model with the time participants took to categorise a word as a fluent access cue as a predictor and the time they needed to generate a memory in response to this word as a measure. Participants were the random factor of the model. We found no effect of the prediction time on AM retrieval fluency, $b = 1.614$, $t(95.114) = 0.878$, $p = .382$. Also, we designed linear mixed models with prediction time as a factor and on memory ratings as measures. We found no effect of prediction time on vividness, ease of generation, specificity, personal significance, valence nor intensity, all $Ps > .24$

TABLE 1.
Means (standard deviations) for measures related to fluent access cues and limited access cues in Study 1 and Study 2.

Study 1	Fluent access cues	Limited access cues
Fluency (s)	5.94 (7.75)	8.70 (6.70)
Vividness	5.26 (1.56)	4.61 (1.64)
Ease of Generation	5.55 (1.51)	4.60 (1.71)
Specificity	3.86 (1.86)	3.20 (1.80)
PersoSig	3.86 (1.86)	3.20 (1.80)
Valence	0.88 (1.76)	0.76 (1.57)
Intensity	1.66 (1.06)	1.40 (1.04)

Study 2		
Fluency (s)	6.03 (5.41)	7.54 (5.85)
Vividness	5.35 (1.70)	4.66 (1.82)
Ease of Generation	5.71 (1.45)	5.02 (1.60)
Specificity	5.33 (1.77)	4.74 (1.81)
PersoSig	3.86 (1.94)	3.36 (1.90)
Valence	0.83 (1.74)	0.96 (1.58)
Intensity	1.63 (1.02)	1.39 (1.02)

Autobiographical Memory Performance. We designed linear mixed models with AM retrieval fluency (in seconds) as predictor and with the memory ratings as measures to assess whether AM generation fluency predicted the memory ratings. Participants were the random factor of the models. We found that a greater fluency (i.e., shorter retrieval latency) led to AMs rated as more vivid, b = -0.059474, t(796.483512) = -6.679, p < .001, more easily generated, b = -0.124963, t(797.787365) = -15.98, p < .001, more specific, b = -0.038255, t(800.440982) = -4.118, p < .001, more personally significant, b = -0.05507, t(796.50148) = -5.398, p < .001 and more emotionally intense, b = -0.017746, t(749.255710) = -2.945, p < .01. However, we found no link between AM generation fluency and AM valence, b = -0.017153, t(711.778367) = -1.791, p = .07.

We calculated repeated-measures correlations between all memory ratings using the *rmcorr* package (version 0.5.2, Bakdash & Marusich, 2017). Overall, most of the correlations were positive, moderate to strong, and significant (Table 2).

TABLE 2.
Correlations between memory ratings (r value).

	Vividness	Ease of generation	Specificity	PersSig
Ease of generation	.61***			
Specificity	.62***	.43***		
PersSig	.36***	.36***	.37***	
Valence	.09**	.08*	.11**	.41***
Intensity	.28***	.26**	.21**	.56***

Significance level: *.05 **.01 ***.001

Cue Word Analysis. We carried out three types of analyses to investigate whether word properties influenced participants' responses during the prediction task. First, we designed linear mixed models with the type of cue as predictor and with the word properties as measures to assess whether fluent access cues and limited access cues selected by each participant would differ in terms of familiarity, concreteness, imageability, and age of acquisition. Participants were the random factor of the models. We found that fluent access cues were less concrete, b = -0.036, $t(6000)$ = -2.486, $p < .05$, and more familiar, b = 0.15, $t(6026)$ = 3.514, $p < .001$, than limited access cues. We found that fluent access cues and limited access cues do not differ in terms of imageability, b = 0.02, $t(6000)$ = 1.6218, p = .106, nor in term of age of acquisition, b = -0.043, $t(6000)$ = -1.062, p = .288.

Second, a negative binomial regression model was designed to assess whether psycholinguistic properties of the words could predict the number of times the words were chosen as fluent access cues. Age of acquisition, familiarity, imageability, and concreteness were the predictors of the model. The number of times a word was chosen as a fluent access cue was the measure. We found that familiarity and imageability were positively associated with the number of times a word was chosen as a fluent access cue (respectively, *IRR* = 1.09, 95% IC [1.02; 1.18], $p <$ .01; *IRR* =1.77, 95% IC [1.30; 2.41], $p < .001$). However, concreteness was negatively associated with the number of times a word was chosen as a fluent access cue, *IRR* =0.58, 95% IC [0.44; 0.76] $p < .001$. We found no effect of age of acquisition, *IRR* =0.99, 95% IC [0.95;

1.04], p = .89. As an example, the word *Birthday* (‘*Anniversaire’*) was the most chosen fluent access cue (51 participants, associated with an imageability of 3.65, a concreteness of 2.9, a familiarity of 3.57 - these ratings were done on 0-5 scales - and an age of acquisition of 6.51) whereas the word *Nail (‘Clou’)* was the least chosen fluent access cue (6 participants, associated with an imageability of 4.8, a concreteness of 4.76, a familiarity of 3, and an age of acquisition of 6.14).

Third, we designed a linear mixed model with word properties as predictors, and with the time participants took to categorise a word as a fluent access cue as a measure. We found that prediction time increased (i.e., slowed) with word concreteness, b = 0.09782, $t(2618)$ = 4.253, $p < .001$, but decreased with word imageability, b = -0.09422, $t(2618)$ = -3.346, $p < .001$, and with word familiarity, b = -0.02142, $t(2618)$ = -3.606, $p < .001$. We found no effect of age of acquisition on prediction time, b = -0.00082, $t(2618)$ = -0.215, p = .83.

Finally, we carried out analyses to investigate whether word properties influenced participants’ responses during the AM task. Thus, we designed a linear mixed model with word properties as predictors and AM generation fluency as measure. We found that no word property predicted AM generation fluency, all *Ps* > .2.

Interim Discussion

Our participants selected words which they judged as being more likely to lead to the generation of an autobiographical memory. These words selected yielded faster retrieval times in a subsequent AM task. Moreover, AMs retrieved in response to fluent access cues were subjectively evaluated as more fluent, vivid, specific, personally significant and emotionally intense. However, even if participants’ metacognitive judgements predicted their responses at the AM task, the speed with which they made their judgements did not predict retrieval fluency nor memory ratings.

This first study showed that people were able to quickly distinguish cue words that lead to fluent access of AM from cue words that would not. That is, participants demonstrate metacognitive awareness of cues pertaining to their AM retrieval. Importantly, 82% of participants indicated that they tried to retrieve memories while categorising the words and that they often (60% of the time) relied on their retrieval to make their decision. Although the use of such a strategy did not seem to influence the results when we added the frequency of using this strategy into our models, it remains possible that the metacognitive evaluation made in the prediction phase is based on a retrieved memory actually coming to mind, and an evaluation of that information rather than any pre-retrieval rapid access to an autobiographical representation or feeling accompanying such processing.

Our response to this possibility was to run the experiment again but with an even shorter response window of one second as opposed to two seconds. To foreshadow the general discussion of the results, however, we point out here that a) two seconds would be too rapid for most accounts of voluntary retrieval and b) stimuli characteristics - especially familiarity - tended to influence the decisions made in our prediction phase without having the same pattern of influence ultimately for the memories retrieved in the final phase of the experiment.

Experiment 2

After analysing Study 1 data, we aimed to replicate these findings and to reduce the high number of retrieval attempts observed during the prediction task. The procedure of Study 2 was the same as in Study 1, except that we diminished the presentation time of the prediction task. With this variation, we tried to diminish the frequency with which participants tried to retrieve AMs during the prediction part, that is, to emphasise the use of a feeling of fluency and to decrease the probability of retrieval[2].

Method

Participants and Material . The words used in this experiment were the same as those used in Study 1, and the task was implemented on the same software and testing environment. Based on Study 1, we aimed to recruit 40 participants. Overall, we collected data from 42 participants (M_{age} = 21.4, SD_{age} = 4.2; 66% of female). Three participants failed the attention checks, leading to a sample size of 39 participants (M_{age} = 21.5, SD_{age} = 4.4; 64% of female)

[2] We also started a third identical experiment, giving participants only 0.5 seconds to make the prediction judgement. We stopped recruiting participants after collecting data from 15 participants. Among these 15 participants, 9 failed the prediction task. That is, they failed to press the spacebar at least eight times so as to have eight words for which they would

Procedure. The four phases of Study 1 were identical to Study 2, except for the prediction task in which words were on the screen for only 1 s.

Data Preprocessing and Analyses. As in Study 1, all thoughts generated during the AM task were read by the first author and a neutral judge. They categorised each description as a specific AM or as something else. They categorised the 30 first descriptions together and then they carried out this task alone and compared their responses. Disagreements were discussed and solved so that all thoughts could be categorised. Agreement rate and Cohen's kappa could not be provided as we lost one judge's original file. Measures (e.g., generation time, vividness rating) for AM responses that were not categorised as specific AMs were removed from the analyses. Out of the 615 recorded descriptions, 541 have been categorised as memories (88%) and 74 have been categorised as something else (12%)

Participants who failed 3 or more attention checks were removed from the analyses (n = 3). Each participant retrieved AMs in response to both words they categorised as fluent access cues and words they categorised as limited access cues, thus, contrary to Study 1, we did not remove participants without AMs for a type of cue. For analyses involving response times, we removed all observations above or below 2.5 SD of the mean of the individual (n = 20).

We carried out the same analyses as in Study 1, using R software (version 4.1.1; R Core Team, 2017). In parallel to the analyses presented below, we also analysed the data using RT as a measure using a log transformation and found that the significance of the results was unchanged.

Results

With only 1 second to make their decision, in total, 29 participants[3] (76% of Study 2's participants compared to 82% in Study 1) indicated that they tried to retrieve AMs during the prediction task. On average, participants indicated that they relied on their AM retrieval 58% of the time (compared to 60% in Study 1). These subjective reports suggest that we were somewhat unsuccessful in reducing what participants described as a retrieval strategy.

---

generate memories in the AM phase. Furthermore, all participants failed the attention checks and indicated that the presentation time was far too short to make a decision.

[3] One participant's response for this question was not been recorded.

Effects of Cue Type on Retrieval Processes. We designed a linear mixed model with participants as a random factor to investigate the effect of the type of cue (i.e., limited access cue or fluent access cue) on AM retrieval latency. The type of cue was used as a predictor and the AM generation time as a measure: $b = -1.4410$, $t(41.5755) = -2.381$, $p < .05$. Overall, AMs were retrieved 1.4 s faster in response to fluent access cues than in response to limited access cues (Table 1), confirming the pattern found in Study 1.

We designed a linear mixed model with participants as a random factor to investigate the effect of the type of cue on the ease of generation rating. The type of cue was used as a predictor and the ease of generation rating as a measure: $b = 0.6781$, $t(505.3121) = 5.575$, $p < .001$. Participants rated the generation of an AM in response to a fluent access cue as easier than in response to a limited access cue, as in Study 1.

Using a linear model with the type of cue and the frequency with which participants relied on their voluntary retrieval to categorise words as fluent access cues or limited access cues as predictors, we found that the effect of the type of cue on retrieval fluency was not modulated by the frequency with which participants relied on this strategy, $b = 0.03413$, $t(27) = 0.688$, $p = .4975$, $PRE = 0.017$. We carried out another analysis including participants who did not try to retrieve AMs and found a similar result: $b = -0.02977$, $t(36) = -1.010$, $p = .319$, $PRE = 0.028$. Similarly, such potential retrieval did not modulate the effect of the type of cue on ease of generation ratings, both when the analysis contained only participants who tried to retrieve AMs, $b = -0.002794$, $t(27) = -0.389$, $p = .70$, $PRE = 0.006$ and when the analysis contained all participants, $b = 0.00079$, $t(36) = 0.019$, $p = .985$, $PRE = 0.000$.

We designed linear mixed models with the type of cue as predictor and with the memory ratings as measures to assess whether fluent access cues would lead to different AMs than limited access cues. Participants were the random factor of the models. We found that AMs retrieved in response to fluent access cues were rated as more vivid, $b = 0.65$, $t(39.9560) = 3.777$, $p < .001$, more specific, $b = 0.5877$, $t(504.6056) = 4.23$, $p < .001$, more important, $b = 0.4759$, $t(505.1844) = 3.13$, $p < .01$ and more emotionally intense, $b = 0.23$, $t(507.79079) = 2.743$, $p <$

.01. However, we found no effect of the type of cue on AM valence, $b = 0.14$, $t(511.77061) = 0.983$, $p = .326$.[4]

We designed a linear mixed model with the time participants took to categorise a word as a fluent access cue as a predictor and the time they needed to retrieve a memory in response to this word as a measure. Participants were the random factor of the model. Also, we designed linear mixed models with prediction time as a factor and on memory ratings as measures. We found no effect of prediction time on AM generation fluency nor on memory ratings, all *Ps* > .18

Autobiographical Memory Performance. We designed linear mixed models with AM generation fluency as predictor and with the memory ratings as measures to assess whether AM generation fluency could predict the memory ratings. Participants were the random factor of the models. We found that a greater fluency led to AMs rated as more vivid, $b = -0.09844$, $t(517.74293) = -7.58$, $p < .001$, more easily retrieved, $b = -0.139827$, $t(509.602034) = -14.55$, $p < .001$, more specific, $b = -0.06365$, $t(518.17221) = -4.769$, $p < .001$, more personally significant, $b = -0.05225$, $t(519.48935) = -4.769$, $p < .001$ and more emotionally intense, $b = -0.022625$, $t(521.605234) = -2.836$, $p < .01$. However, we found no relationship between AM generation fluency and AM valence, $b = -0.006742$, $t(504.566785) = -0.508$, $p = .612$.

We calculated repeated-measures correlations between all memory ratings using the *rmcorr* package (version 0.5.2, Bakdash & Marusich, 2017). Overall, most of the correlations were moderate to strong and significant (Table 3).

---

[4]Based on a reviewer's suggestion, we analysed the effect of the type of cue on memory valence using a categorical approach. For each of our studies, we designed binomial generalised linear mixed models to inspect whether the type of cue would predict the valence of the memory. Memories associated with a negative rating (i.e., < 0) were categorised as negative, those associated with a positive rating (i.e., > 0) were categorised as positive. Those memories with a neutral rating (i.e., 0) were removed from the analysis (n= 163 (20%) in study 1, n= 94 (17%) in study 2). We found no effect of the type of cue on memory valence in both studies (respectively b = -0.1368, p = .481 in study 1; b = 0.1223, p = .566 in study 2).

TABLE 3.
Correlations between memory ratings (r values).

	Vividness	Ease of generation	Specificity	PersSig
Ease of generation	0.64***			
Specificity	0.63***	0.43***		
PersSig	0.40***	0.37***	0.34***	
Valence	0.11***	0.12**	0.10*	0.44***
Intensity	0.32***	0.29***	0.26***	0.58***

Significance level: *.05 **.01 ***.001

Cue Word Analysis. As in Study 1, three types of analyses aimed to investigate whether word properties could have influenced participants' responses during the prediction part. First, We designed linear mixed models with the type of cue as predictor and with the word properties as measures to assess whether fluent access cues and limited access cues selected by each participant would differ in terms of familiarity, concreteness, imageability, and age of acquisition. Participants were the random factor of the models. We found that fluent access cues were less concrete, b = -0.005987, t(3899) = -3.273, p < .01, and more familiar, b = 0.2435, t(42.01)= 5.797, p < .001, than limited access cues. We found that fluent access cues and limited access cues do not differ in terms of imageability, b = 0.02464, t(3899) = 1.507, p = .132, nor in term of age of acquisition, b = -0.02461, t(3899.)= -0.473, p = .637.

Second, a negative binomial regression model was designed to assess whether psycholinguistic properties of the words could predict the number of times the words have been chosen as fluent access cues. Age of acquisition, familiarity, imageability, and concreteness were the predictors of the model. The number of times a word was chosen as a fluent access cue was the measure. We found that familiarity and imageability were positively associated with the number of times a word was chosen as a fluent access cue (respectively, *IRR* = 1.18, 95% IC [1.10; 1.28], p < .001; *IRR* =2.98, 95% IC [2.05; 4.43], p < .001). However, concreteness was negatively associated with the number of times a word was chosen as a fluent access cue, *IRR* =0.37, 95% IC [0.27; 0.50], p < .001. We found no effect of age of acquisition, *IRR* =1, 95% IC [0.95; 1.05]; p = .91. Interestingly, in this experiment, *Family* ('*Famille*') was the most chosen fluent access cue (35 participants, associated with an imageability of 4.24, a concreteness of 3.56, a familiarity of 4.69 -these ratings were done on 0-5 scales- and an age of acquisition of 6.42) and *Acacia* was the least chosen fluent access cue (1 participant, associated with an imageability of 2.44, a concreteness of 3.85, a familiarity of 4.38, and an age of acquisition of 3.67).

Third, we designed a linear mixed model with word properties as predictors, and with the time participants took to categorise a word as a fluent access cue as a measure. We found that prediction time increased with word concreteness, $b = 0.05416$, $t(1369) = 3.590$, $p < .001$, but decreased with word imageability, $b = -0.07124$, $t(1369) = -3.617$, $p < .001$, and with word familiarity, $b = -0.01383$, $t(1372) = -3.966$, $p < .001$. We found no effect of age of acquisition on prediction time, $b = 0.001884$, $t(1366) = 0.825$, $p = .409$.

Finally, we carried out analyses to investigate whether word properties influenced participants' responses during the AM task. We designed a linear mixed model with word properties as predictors and AM generation fluency as measure. We found that word familiarity was the only factor predicting AM generation fluency, $b = 0.5704$, $t(504.4207) = 2.243$, $p < .05$

Interim Discussion

Study 2 replicated all the cue type effects found in Study 1. With only 1 s to make their metacognitive judgement, participants were still able to correctly categorise words as fluent access cues or limited access cues. As in Study 1, a majority of participants reported trying to retrieve memories during the prediction task, and indicated that they tended to rely on their AM retrieval to categorise words as fluent access or limited access cues (See Table 4 for details about participants' predictions). Again, the use of this strategy had no influence on observed monitoring abilities. Fluent access cues were again associated with a better AM generation fluency, and higher memory ratings (Table 1).

Relationships found in Study 1 between AM generation fluency and memory ratings, and between memory ratings were replicated in Study 2. Similarly, cue word properties influenced participants' judgments in the prediction phase the same way they influenced it in Study 1. One notable difference between our two studies is the influence of cue word familiarity on AM generation fluency found in Study 2. This result contradicts Study 1 and other studies (e.g., Williams et al., 1999). Otherwise, similar but not identical influences of word properties were found in our two studies.

TABLE 4. Participants' responses in the prediction task.

Study 1		
	Mean prediction time (SD) in sec	1.18 (0.33)
	Mean amount of words categorised as fluent access cues (SD; range)	43 (17; 9-80)
	Mean frequency with which participants relied on a voluntary retrieval to make their predictions (SD)	60 (24)
Study 2		
	Mean prediction time (SD) in sec	0.69 (0.14)
	Mean amount of words categorised as fluent access cues (SD; range)	33 (20; 11-74)
	Mean frequency with which participants relied on a voluntary retrieval to make their predictions (SD)	58 (24)

General discussion

Multiple control processes are supposed to occur during AM retrieval. Mostly, these processes have been theorised in relation to AM voluntary retrieval (Barzykowski & Moulin, 2022; Conway, 2005). Paradoxically, most of these have been described as non-voluntary (sometimes the term "non-conscious" has been used too, e.g., Conway et al., 2004). In this article, we aimed to explore whether individuals could voluntarily monitor their AM retrieval. For that purpose, we designed a new task assessing metacognitive abilities related to the earliest stage of AM retrieval. This task allowed us to investigate whether people could predict which cues would facilitate or limit access to autobiographical memories. Moreover, it allowed us to investigate which stimuli properties influence such metacognitive judgements related to AM, and AM retrieval.

In two studies, participants correctly distinguished cue words which facilitated their AM retrieval from cue words that did not, in keeping with results using a different paradigm (Matsumoto, 2022). Two key analyses point to this result. The first was based on an objective measure of ease of AM retrieval, the retrieval fluency, illustrated by response times. The second analysis was based on a subjective evaluation, made by participants on a 1-7 scale. They rated how easy it was to retrieve the memory. Importantly we should note that our reaction times were much shorter than in Matsumoto (2022) where the time to make each evaluation was about 4.5 secs[5] , which raises questions about how the fluency judgements in our task are actually made, which we discuss below, since it seems much more rapid than the time taken either to retrieve a memory or to make an explicit metacognitive evaluation.

Before attempting to explain these results, we first seek to describe the finding in an ecological context. Our data shows that well before the participant is able to report a coherent, fully formed and specific autobiographical memory, they are able to make a well-informed judgement about how likely each cue is to lead to such a memory. Firstly, our results pertain to voluntary retrieval only; we think they apply to instances where Conway's concept of the retrieval mode has been activated. We suggest that when people are in such a retrieval mode they are vigilant to cues in the environment which may aid their autobiographical retrieval. As an example in the real world, at a family reunion, the contents of a shared conversation based on reminiscence will activate representations based on the cues in the conversation, and we propose here that during such a conversation, we will have access to those aspects of the exchanges that would be most likely to lead to vivid and specific memories which we may choose to share with the group. Another example would be in eyewitness situations, where a series of questions around a set of cues or reminders would lead to patterns of activation of memory, and quickly, on-line, the eyewitness will be able to judge the utility of each cue in answering key questions. An

[5] We are grateful to the author for making this mean RT available to us.

interesting follow up would be to consider our question in the context of a breakdown in autobiographical memory function.

We discuss our findings with reference to possible mechanisms which may be at play in this metacognitive awareness of the earliest phase of autobiographical memory retrieval. We suggest that there are several possible interpretations of our data, and these somewhat overlap with Koriat's (2007) description of metacognitive cues. The first possibility is described as being based on intrinsic cues: independent of a retrieval attempt, people evaluate key characteristics of the cue which are more or less predictive of their subsequent memory performance. That is, without necessarily thinking of an autobiographical memory, people generate a model of retrieval with the notion that highly familiar or imageable words are more likely to cue autobiographical memories. In support, we found that a number of characteristics were related to the frequency with which a word was chosen as a fluent access cue, although we must note that when fluent and limited access cues are directly compared, these relationships do not always hold out. However, regardless of the Study or the type of analysis, we consistently found patterns whereby more familiar and less concrete cue words were more likely to be chosen as fluent access cues.

We thus found that word familiarity was determinant for participants' cue word categorisation (and it also facilitated AM retrieval in Study 2). Firstly, these results are coherent with studies in metacognition, which shows that familiarity is a cue to upcoming retrieval success in cue-target episodic word pair tasks (and also in general knowledge) in the FOK paradigm described in the introduction. The FOK literature emphasises the importance of the feeling of familiarity elicited by the cue in FOK judgements, with a greater feeling of familiarity leading to a greater feeling of retrievability (Castel & Middlebrooks, 2016). Mostly, the evidence for such a cue familiarity mechanism in metacognition derives from experimentally-induced familiarity (e.g., Hanczakowski et al., 2013; Metcalfe et al., 1993; Reder, 1987) rather than the stimuli characteristic of familiarity as used here. Our proposal is that as well as experimentally induced familiarity for recently encountered information, people are also sensitive to lifetime exposure to concepts, words and ideas, and this form of familiarity influences metacognitive evaluations. In support, a series of studies conducted by Mendes & Undorf (2022) demonstrated that familiar cues (highly frequent words) were judged as more memorable than no familiar cues (rare words).

Thus we propose that cue word familiarity could have been an intrinsic cue leading to a greater feeling of fluency (and feeling of familiarity) during the prediction phase. This intrinsic evaluation of future performance is predictive because in fact, greater cue familiarity leads to faster AM retrieval (Gurguryan et al., 2023; Robin & Moscovitch 2014). Gurguryan et al. (in pre-print) showed that cue words associated with a high lifetime exposure (i.e., a high familiarity, e.g., the cue word *bottle*, in opposition to low familiarity, e.g., *willow*) facilitated AM retrieval.

The fact that word concreteness was also determinant in categorising words as limited access cues and that high word concreteness was associated with lower ease of generation ratings in Study 1 was not expected. AM studies using cue words with different levels of concreteness (e.g., Rubin & Schulkind 1997), and studies comparing AM generation in response to cue words and in response to real (concrete) objects (e.g., Kirk & Berntsen 2018) found that more concrete cues better facilitate AM retrieval. We are not sure why participants responded by judging less concrete words as more likely to lead to autobiographical retrieval, but it is likely to be due to the complex evaluation of multiple characteristics of the cue words by our participants, and the fact that our cue words were all nouns and varied across different characteristics in an inter-related manner. *Birthday* seems intrinsically like a word likely to cue autobiographical memories, but is a much less concrete word than *Acacia*, as an example. Finally, imageable words are supposed to facilitate AM retrieval (e.g., Williams et al., 1999) and word imageability was associated with fluent access cues in both of our studies. However, imageability had no effect on AM retrieval during the AM task, although it was related to a few of our subjective memory ratings.

A second mechanism by which people might make these metacognitive judgements is mnemonic, and based on access to information (Koriat, 2008). This explanation stresses the involvement of monitoring in AM retrieval. We propose that in the retrieval mode, participants were able to judge how efficient the cue was in the pre-retrieval stage. Participants are capable as such of monitoring their AM function directly rather than just evaluate the stimulus characteristics of the cue word and have either responded based on the amount of activation target memories received or based on the amount of associated material which was accessible for retrieval.

We propose that this mnemonic access takes two forms. First, it may simply address the fluency of processing the cue word and the subsequent retrieval event. Barzykoswki and Moulin (2022) propose that epistemic feelings such as familiarity guide AM retrieval and are particularly important in the pre-retrieval stage in which individuals enter Conway's retrieval mode. They stress that such feelings arrive before content, and that these feelings are in fact content-free. According to this idea, in our procedure, a feeling of fluency while being in the retrieval mode may have been interpreted as a sensation of memory accessibility, or inaccessibility, influencing the categorisations of words. In fact, aside from the very short response times, with a cut off of 1 s in the second study, which makes retrieval unlikely, we do not know if any content came to mind or not during the prediction phase. As such it is difficult to support this mechanism given the current design and data. However, the notion of a threshold was supported since memories retrieved in response to fluent access cues were rated as more vivid, specific, personally significant, and emotionally intense than those retrieved in response to limited access cues. In the context of the threshold hypothesis, AMs that are the most vivid, emotionally intense, important, and unique are more easily accessed as those factors mean that the memory is more likely to pass the threshold (Barzykowski et al., 2019; Barzykowski & Staugaard, 2016, 2018). Also, as stated by the threshold hypothesis, we found that the more easily memories were retrieved (i.e., illustrated by retrieval latencies), the more they were vivid, specific, personally significant, and emotionally intense (Barzykowski et al., 2019).

The second aspect of monitoring, in contrast, highlights the possible content, often described as partial information, which comes to mind during a retrieval attempt and which can be used to make accurate FOK judgements in episodic cue-target tasks. This corresponds to the second stage of FOK judgements of the interactive hypothesis described earlier (Koriat & Levy-Sadot, 2001). A majority of participants in our experiment indicated that they tried to retrieve memories during the prediction task and that they relied on their retrieval for a majority of their responses. If they indeed retrieved complete AMs, they could have carried out a task in which they have retrieved the same memories twice, first during the prediction (leading to responses and thus the categorisation of the current cue as a fluent access cue) and then during the AM task, explaining why AM retrieval was as easier in response to fluent access cues. However, going against this explanation, we found twice that the frequency with which participants relied on their AM retrieval did not modulate the difference of ease of AM generation between fluent access cues and limited access cues. Furthermore, as participants needed more time than was given during the prediction task to retrieve AMs in response to fluent access cues (i.e., approximately 6 s in both studies), and in line with the mean AM retrieval time observed in the literature, it seems unlikely that they managed to retrieve many complete memories during the prediction task. More likely however, they may have retrieved partial information about their memories. Aligning with the interactive hypothesis, they may have based their predictions by monitoring this partial accessible information. What needs to be investigated is whether they have categorised a cue as a fluent access cue because partial information was accessed, meaning that limited access cues gave access to no information during the prediction task, or because the amount of information provided by certain cues reached a decision threshold. A way to address this question would be to control whether each cue word gives access to partial information or to evaluate the amount of information accessed during cue presentation.

The presence of metacognition during early AM retrieval adds to the literature that investigates metacognition later in retrieval (e.g. Mazzoni and Kirsch, 2002). Such early-stage metacognitive evaluations can be used to determine whether a retrieval attempt should take place. Mazzoni and Hanczakowski (2011) proposed that an evaluation of event plausibility influences whether a retrieval is attempted and how much effort should be allocated for this retrieval. Autobiographical memory beliefs drive the efforts put into the retrieval attempt, and are influenced by the event's plausibility, the presence of recollection, and social feedback (Scoboria and Henkel, 2020). Our proposal here is that the metacognitive access that we have demonstrated here is a process that presumably acts to assist in the elaboration of cues. It seems we are able to rapidly identify cues as potentially useful for retrieving autobiographical content, and so this would presumably help us to guide retrieval processes towards relevant material. Conway's example (Conway et al., 2001; p.495) is of a person given the cue word *Beach*. They ' … initially elaborate this into a general event cue such as "*Where did we go on holiday last year?"* And access the information *"France"* and then use this knowledge to initiate a further search cycle and so on until knowledge is activated that satisfies the constraints, or some subset of these, of the centrally generated retrieval verification model.'

Conclusion

Conway (e.g., Conway et al., 2001) proposed that voluntary autobiographical retrieval involved iterative *search–evaluate–elaborate* cycles. The central part of his proposition is the capacity to evaluate the success and quality of the partially retrieved information on the way to retrieving a fully formed autobiographical representation. To date, evidence for this evaluation of autobiographical memory was lacking. In two studies proposing an original design we showed that participants were able to predict which words would help them to retrieve a specific AM and which words would be less likely to help. Based on Barzykowski & Mazzoni (2021) proposal of the sequential phases of retrieval, we argue that our participants were able to monitor their AM pre-retrieval stage, that is, the stage in which they enter the retrieval mode[6]. This apparently rapid, and arguably relatively automatic process may rely on epistemic feelings and retrieval of partial information to shape AM retrieval constituting a metacognitive access to autobiographical memory at the earliest stages of retrieval (Barzykowski & Moulin, 2022; Moulin et al., 2022). We have identified a paradigm therefore which may be useful in a neuroimaging context for elucidating the neural mechanisms in autobiographical memory construction and in developing Conway's notion that frontal networks control and monitor the construction of autobiographical memories.

Finally, we note that participants reported that they relied on retrieval to make their predictions. Interestingly, asking for a similar prediction, Matsumoto (2022) indicated that participants rarely recalled memories to make their judgement. We do not think that participants actually managed to access fully-formed memories during the prediction task. However, a priority would be to develop designs with more control over a potential AM generation during the prediction task, in order to further isolate metacognitive processes from a potential AM retrieval during the prediction task.

[6] In one of his latter keynote presentations in Grenoble, France, Conway (2017; a similar argument has been developed in Conway & Loveday, 2015) criticised the introspective approach towards studying autobiographical memory. He again stated that '... the system elaborates a cue in some iterative manner until a specific and effective cue is generated, enters the retrieval process and knowledge in long-term memory is then accessed.' He added that 'It is important to note that this 'generative process' does not happened [sic] consciously. Sometimes aspects of it may enter consciousness, sometimes it may be completely non-conscious.' and 'I could go on but the point is plain and that is that we cannot introspect on non-conscious memory retrieval process although we can be aware of some of their outputs.' Here we develop this statement: our participants may feel like they have initiated retrieval, but all that is known to them is what fragments may come to mind or how fluently the cue has been processed.

Data availability Statement

These studies were pre-registered in the Open Science Framework. All data, experiment scripts, and analysis scripts are available online (https://osf.io/r2j74/?view_only=d19767270c294be1b50b0c3bb53db43c).

Disclosure of interest

The authors report no conflict of interest

References

Arango-Muñoz, S. (2011). Two Levels of Metacognition. *Philosophia*, *39*(1), 71-82. https://doi.org/10.1007/s11406-010-9279-0

Arango-Muñoz, S. (2014). The nature of epistemic feelings. *Philosophical Psychology*, *27*(2), 193-211. https://doi.org/10.1080/09515089.2012.732002

Bakdash, J. Z., & Marusich, L. R. (2017). Repeated Measures Correlation. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00456

Barzykowski, K., & Mazzoni, G. (2021). Do intuitive ideas of the qualities that should characterize involuntary and voluntary memories affect their classification? *Psychological Research*, *86*(1), 170-195. https://doi.org/10.1007/s00426-020-01465-3

Barzykowski, K., & Moulin, C. (2022). Are involuntary autobiographical memory and déjà vu natural products of memory retrieval? *The Behavioral and Brain Sciences*, 1-66. https://doi.org/10.1017/S0140525X22002035

Barzykowski, K., Niedźwieńska, A., & Mazzoni, G. (2019). How intention to retrieve a memory and expectation that a memory will come to mind influence the retrieval of autobiographical memories. *Consciousness and Cognition: An International Journal*, *72*, 31-48. https://doi.org/10.1016/j.concog.2019.03.011

Barzykowski, K., & Staugaard, S. R. (2016). Does retrieval intentionality really matter? Similarities and differences between involuntary memories and directly and generatively retrieved voluntary memories. *British Journal of Psychology*, *107*(3), 519-536.

Barzykowski, K., & Staugaard, S. R. (2018). How intention and monitoring your thoughts influence characteristics of autobiographical memories. *British Journal of Psychology*, *109*(2), 321-340. https://doi.org/10.1111/bjop.12259

Bastin, C., Besson, G., Simon, J., Delhaye, E., Geurten, M., Willems, S., & Salmon, E. (2019). An integrative memory model of recollection and familiarity to understand memory deficits. *Behavioral and Brain Sciences*, *42*(e281), 1-60. https://doi.org/10.1017/S0140525X19000621

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). *Parsimonious mixed models* (arXiv Preprint arXiv: 1506.04967.). https://doi.org/10.48550/arXiv.1506.04967

Carreras, F., & Moulin, C. J. A. (2023). Evidence for a metacognitive awareness of autobiographical memory organisation. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-34389-0

Castel, A. D., & Middlebrooks, C. D. (2016). Monitoring Memory in Old Age : Impaired, Spared, and Aware. In J. Dunlosky, S. U. K. Tauber, & S. McGillivray (Éds.), *The Oxford handbook of metamemory* (p. 519-536). Oxford University Press.

Conway, M. A. (2005). Memory and the self. *Journal of Memory and Language*, *53*(4), 594-628. https://doi.org/10.1016/j.jml.2005.08.005

Conway, M. A., & Bekerian, D. A. (1987). Organization in autobiographical memory. *Memory & Cognition*, *15*(2), 119-132. https://doi.org/10.3758/BF03197023

Conway, M. A., Justice, L. V., & D'Argembeau, A. (2019). The Self-Memory System Revisited : Past, Present, and Future. In J. Mace (Éd.), *The organization and structure of autobiographical memory* (pp. 28–51). Oxford University Press. https://doi.org/10.1093/oso/9780198784845.003.0003

Conway, M. A., & Loveday, C. (2015). Remembering, imagining, false memories & personal meanings. *Consciousness and Cognition*, *33*, 574-581. https://doi.org/10.1016/j.concog.2014.12.002

Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review*, *107*(2), 261-288. https://doi.org/10.1037/0033-295X.107.2.261

Conway, M. A., Pleydell-Pearce, C. W., & Whitecross, S. E. (2001). The Neuroanatomy of Autobiographical Memory : A Slow Cortical Potential Study of Autobiographical Memory Retrieval☆. *Journal of Memory and Language*, *45*(3), 493-524. https://doi.org/10.1006/jmla.2001.2781

Conway, M. A., & Loveday, C. (2015). Remembering, imagining, false memories & personal meanings. *Consciousness and cognition*, 33, 574-581. https://doi.org/10.1016/j.concog.2014.12.002

Conway, M. A., Singer, J. A., & Tagini, A. (2004). The Self and Autobiographical Memory : Correspondence and Coherence. *Social Cognition*, *22*(5), 491-529. https://doi.org/10.1521/soco.22.5.491.50768

Conway, M. A., Turk, D. J., Miller, S. L., Logan, J., Nebes, R. D., Meltzer, C. C., & Becker, J. T. (1999). A Positron Emission Tomography (PET) Study of Autobiographical Memory Retrieval. *Memory*, *7*(5-6), 679-703. https://doi.org/10.1080/096582199387805

Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience*, *8*. https://www.frontiersin.org/article/10.3389/fnhum.2014.00443

Green, P., & MacLeod, C. J. (2016). SIMR : an R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, *7*(4), 493-498. https://doi.org/10.1111/2041-210X.12504

Gurguryan, L., Yang, H., Köhler, S., & Sheldon, S. (2023). *Characterizing the contributions of cue familiarity for the retrieval of autobiographical memories* [Preprint]. In Review. https://doi.org/10.21203/rs.3.rs-3157762/v1

Hanczakowski, M., Pasek, T., Zawadzka, K., & Mazzoni, G. (2013). Cue familiarity and 'don't know' responding in episodic memory tasks. *Journal of Memory and Language*, *69*(3), 368-383. https://doi.org/10.1016/j.jml.2013.04.005

Kirk, M., & Berntsen, D. (2018). A short cut to the past : Cueing via concrete objects improves autobiographical memory retrieval in Alzheimer's disease patients. *Neuropsychologia*, *110*, 113-122. https://doi.org/10.1016/j.neuropsychologia.2017.06.034

Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review, 100*(4), 609-639. https://doi.org/10.1037/0033-295X.100.4.609

Koriat, A. (2007). Metacognition and consciousness. In *The Cambridge handbook of consciousness* (p. 289-325). Cambridge University Press. https://doi.org/10.1017/CBO9780511816789.012

Koriat, A. (2008). Subjective confidence in one's answers : The consensuality principle. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 945-959. https://doi.org/10.1037/0278-7393.34.4.945

Koriat, A., & Levy-Sadot, R. (2001). The combined contributions of the cue-familiarity and accessibility heuristics to feelings of knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(1), 34-53. https://doi.org/10.1037/0278-7393.27.1.34

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package : Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*, 1-26. https://doi.org/10.18637/jss.v082.i13

Mace, J. H., Staley, B. J. A., & Sopoci, M. K. (2021). When trying to recall our past, all roads lead to Rome : More evidence for the multi-process retrieval theory of autobiographical memory. *Memory & Cognition*, *49*(3), 438-450. https://doi.org/10.3758/s13421-020-01099-1

Matsumoto, N. (2022). *Meta-memory (prediction) of specific autobiographical recall : An experimental approach using a modified Autobiographical Memory Test* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/gvwft

Mazzoni, G., & Hanczakowski, M. (2011). Metacognitive Processes before and during Retrieval. In P. Higham & J. Leboe (Éds.), *Constructions of Remembering and Metacognition : Essays in Honour of Bruce Whittlesea* (p. 91-106). Palgrave Macmillan. https://link.springer.com/book/10.1057/9780230305281

Mazzoni, G., & Kirsch, I. (2002). Autobiographical memories and beliefs : A preliminary metacognitive model. In T. J. Perfect & B. L. Schwartz (Éds.), *Applied metacognition* (p. 121-145). Cambridge University Press. https://doi.org/10.1017/CBO9780511489976.007

Mendes, P. S., & Undorf, M. (2022). On the pervasive effect of word frequency in metamemory. *Quarterly Journal of Experimental Psychology*, *75*(8), 1411-1427. https://doi.org/10.1177/17470218211053329

Metcalfe, J., Schwartz, B. L., & Joaquim, S. G. (1993). The cue-familiarity heuristic in metacognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(4), 851-861. https://doi.org/10.1037/0278-7393.19.4.851

Miceli, A., Wauthia, E., Lefebvre, L., Ris, L., & Simoes Loureiro, I. (2021). Perceptual and Interoceptive Strength Norms for 270 French Words. *Frontiers in Psychology*, *12*. https://www.frontiersin.org/articles/10.3389/fpsyg.2021.667271

Michaelian, K., & Arango-Muñoz, S. (2014). Epistemic Feelings, Epistemic Emotions : Review and Introduction to the Focus Section. *Philosophical Inquiries*, *2*(1), Article 1. https://doi.org/10.4454/philinq.v2i1.79

Morris, R. G., & Mograbi, D. C. (2013). Anosognosia, autobiographical memory and self knowledge in Alzheimer’s disease. *Cortex*, *49*(6), 1553-1565. https://doi.org/10.1016/j.cortex.2012.09.006

Morson, S. M., Moulin, C. J. A., & Souchay, C. (2015). Selective deficits in episodic feeling of knowing in ageing : A novel use of the general knowledge task. *Acta Psychologica*, *157*, 85-92. https://doi.org/10.1016/j.actpsy.2015.02.014

Moulin, C. J. A., Carreras, F., & Barzykowski, K. (2022). The phenomenology of autobiographical retrieval. *WIREs Cognitive Science*, e1638. https://doi.org/10.1002/wcs.1638

Nelson, T., & Narens, L. (1990). Metamemory : A Theoretical Framework and New Findings. In G. H. Bower (Éd.), *Psychology of Learning and Motivation* (Vol. 26, p. 125-173). New York: Acedemic Press.

Nelson, T., & Narens, L. (1994). Why investigate metacognition. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (Vol. 13, pp. 1–25). MIT Press

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2 : Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195-203. https://doi.org/10.3758/s13428-018-01193-y

Proust, J. (2007). Metacognition and metarepresentation : Is a self-directed theory of mind a precondition for metacognition? *Synthese*, *159*(2), 271-295. https://doi.org/10.1007/s11229-007-9208-3

Reder, L. M. (1987). Strategy selection in question answering. *Cognitive Psychology*, *19*(1), 90-138. https://doi.org/10.1016/0010-0285(87)90005-3

Reder, L. M., & Ritter, F. E. (1992). What determines initial feeling of knowing ? Familiarity with question terms, not with the answer. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(3), 435-451. https://doi.org/doi/10.1037/0278-7393.18.3.435

Robin, J., & Moscovitch, M. (2014). The effects of spatial contextual familiarity on remembered scenes, episodic memories, and imagined future events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(2), 459-475. https://doi.org/10.1037/a0034886

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, *104*(3), 192-233. https://doi.org/10.1037/0096-3445.104.3.192

Rubin, D. C., & Schulkind, M. D. (1997). Distribution of Important and Word-Cued Autobiographical Memories in 20-, 35-, and 70-Year-Old Adults. *Psychology and Aging*, *12*(3), 524-535. https://doi.org/10.1037//0882-7974.12.3.524

Scoboria, A., & Henkel, L. (2020). Defending or relinquishing belief in occurrence for remembered events that are challenged : A social-cognitive model. *Applied Cognitive Psychology*, *34*(6), 1243-1252. https://doi.org/10.1002/acp.3713

Uzer, T., Lee, P. J., & Brown, N. R. (2012). On the prevalence of directly retrieved autobiographical memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(5), 1296-1308. https://doi.org/10.1037/a0028142

Welch-Ross, M. K. (1995). An integrative Model of the Development of Autobiographical Memory. *Developmental Review*, *15*, 338-368.

Williams, J. M. G., Healy, H. G., & Ellis, N. C. (1999). The Effect of Imageability and Predicability of Cues in Autobiographical Memory. *The Quarterly Journal of Experimental Psychology Section A*, *52*(3), 555-579. https://doi.org/10.1080/713755828