NEUROSCIENCE of Consciousness

Neuroscience of Consciousness, 2025, **2025(1)**, niaf003
DOI: https://doi.org/10.1093/nc/niaf003
Research Article

On the complexity of metacognitive judgments of memory: evidence from retrospective confidence, feeling of knowing, and older adults

Lucile Meunier-Duperray[1,2,*], Audrey Mazancieux[3], Céline Souchay[1], Christine Bastin[4], Lucie Angel[2], Chris J.A. Moulin[1]

[1]Laboratoire de Psychologie et NeuroCognition (LPNC), CNRS UMR 5105,Université Grenoble-Alpes (UGA), Université Savoie Mont-Blanc (USMB), 1251 rue des Universités, Grenoble 38058, France
[2]Centre de Recherches sur la Cognition et l'Apprentissage (CeRCA), UMR CNRS 7295,Université de Tours, Université de Poitiers, 3 Rue des Tanneurs, Tours 37041, France
[3]Center for Research in Cognition and Neurosciences (CRCN), Université Libre de Bruxelles (ULB), 50 avenue F.D. Roosevelt, Brussels CP191 B–1050, Belgium
[4]GIGA Research CRC Human Imaging, University of Liège, Allée du 6 Août, 8, Liège 4000, Belgium

*Corresponding author. Laboratoire de Psychologie et NeuroCognition (LPNC CNRS 5105), Université Grenoble Alpes, 1251 Avenue Centrale, Domaine Universitaire de St Martin d'Hères, Grenoble 38040, France. E-mail: Lucile.Meunier-Duperray@univ-grenoble-alpes.fr

Abstract

Dissociations in types of memory tasks emerge when comparing feeling-of-knowing (FOK) judgments, predictions of upcoming performance, and retrospective confidence. This pattern has been used to construct theories of metacognitive access to memory, particularly in memory-impaired groups. In particular, older adults' metacognitive sensitivity appears to vary between episodic (impaired) and semantic (intact) memory. However, this could be explained by the limitations of metacognitive measures and/or memory differences. We aimed to test these dissociations of metacognition with aging by comparing metacognitive efficiency in episodic and semantic tasks using two types of judgment: retrospective confidence judgments (RCJs) and FOK judgments. Metacognitive efficiency was estimated in 240 participants aged 19–79 years using a hierarchical Bayesian framework. Results showed that metacognitive efficiency for RCJs declined with age in the semantic task, even though task performance increased with age, while metacognitive efficiency was stable in the episodic task. Surprisingly, metacognitive efficiency was very low (although significantly higher than zero) for both FOK tasks regardless of age compared to similar previous studies. We suggested this might be due to the online testing. These results point to metacognition being multifaceted and varying according to judgment, domains, and populations.

Keywords: confidence judgments; episodic memory; feeling of knowing; healthy aging; metamemory; semantic memory

Introduction

Metamemory, the ability to control and monitor memory (Flavell 1971, Nelson and Narens 1990), is important for adapting cognitive resources and memory strategies (Guerrero Sastoque et al. 2019). It is typically assessed using different types of metacognitive judgments which are self-evaluations about an ongoing memory task. These judgments vary in terms of when the evaluation is made (before or after a task that is prospective or retrospective; Perfect and Stollery 1993, Irak et al. 2023) or in terms of the memory domains and tasks that are targeted (e.g., episodic memory, semantic memory, and short-term memory; Mazancieux et al. 2020a, Lund et al. 2023, McWilliams et al. 2023). By comparing these types of judgment across different domains, researchers can widely explore common mechanisms used in metacognition (Mazancieux et al. 2023). In the current study, we focus on two types of judgment: feeling-of-knowing (FOK) and retrospective confidence judgments (RCJs) because both have been largely used in the literature. Previous studies have revealed a typical pattern in which semantic metacognitive judgments tend to be preserved when semantic memory is altered, but episodic memory judgments are impaired as a function of memory failure (Souchay et al. 2007). Here, we tested metamemory in episodic and semantic tasks using two types of judgment within the same task: FOK judgments and RCJs.

The FOK paradigm (Hart 1965) begins with a cued-recall phase in which participants are asked to recall episodic or semantic information based on a cue. In episodic tasks, cues are typically words associated with a target during an encoding phase (Hicks and Marsh 2002, Perrotin et al. 2006, Souchay et al. 2007, Thomas et al. 2011, Sacher et al. 2013, Morson et al. 2015, Mazancieux et al. 2020a, Bastin et al. 2021, Irak et al. 2023). In semantic tasks, general knowledge questions (Hart 1965, Lachman et al. 1979, Kikyo et al. 2002, Souchay et al. 2007, Elman et al. 2012, Morson et al. 2015, Bastin et al. 2021, Irak et al. 2023) or word definitions are

Received 27 August 2024; Revised 24 December 2024; Accepted 10 February 2025

used (Allen-Burge and Storandt 2000, Mazancieux et al. 2020a). In both tasks and after each trial of the cued-recall phase (or at least after each unrecalled trial), participants are asked to evaluate how confident they are in recognizing the target among several alternatives (i.e., the FOK judgment). Finally, in a recognition phase, participants have to recognize the target associated with each cue between one or more distractors. Because FOK judgments are made before the task to which it is compared (the recognition task), they are prospective judgments.

In contrast, RCJs concern previous recognition tasks and are thus retrospective judgments. During RCJs, participants have to evaluate how confident they are that they give the correct answer in the recognition phase (Lachman et al. 1979, Costermans et al. 1992, Marquié and Huet 2000, Dodson et al. 2007, Chua et al. 2009, Palmer et al. 2014, Voskuilen et al. 2018, Mazancieux et al. 2020a, 2020b, Hertzog et al. 2021, Irak et al. 2023, McWilliams et al. 2023).

In the metacognitive literature, two indices are calculated from these judgments: metacognitive sensitivity and metacognitive bias. Metacognitive sensitivity refers to the ability to discriminate between correct and incorrect responses by linking memory performance (first-order performance) to metamemory judgments (second-order performance) (Fleming and Lau 2014). Metacognitive bias refers to the tendency to give high or low confidence irrespective of first-order performance. This can be computed either using the mean magnitude of metacognitive judgments (Sacher et al. 2015, Mazancieux et al. 2020b, Lund et al. 2023, McWilliams et al. 2023) or using the difference between mean confidence and task performance (Kavé and Halamish 2015, Mazancieux et al. 2020a). Here, we investigate both of these measures, with a particular attention to metacognitive sensitivity.

For metacognitive sensitivity, a typical pattern of impaired episodic FOK and preserved semantic FOK is often observed in older adults (see Devaluez et al. (2023) for a meta-analysis). This dissociation has been also observed in neurological population (Pannu and Kaszniak 2005) and is supported by the idea that episodic and semantic FOK are associated with different brain regions such as greater posterior regions for episodic FOK (Reggev et al. 2011, Elman et al. 2012). These regions are known to be involved in the retrieval process, suggesting that episodic memory and episodic metamemory share common processes (Reggev et al. 2011). Episodic FOK deficits have indeed been found in populations that also have episodic memory deficits such as older adults. Age-related changes in memory differ between episodic memory that declines in aging and semantic memory that remains stable or even increases with age (Mitchell 1989, Nyberg et al. 1996, Jarjat et al. 2021). Given the age-related changes in memory, one of the main issues in studying metacognition in healthy aging is to dissociate the influence of memory processes from a pure metacognitive deficit (Perfect and Stollery 1993). For instance, it has been shown that metacognitive deficits in the clinical population are driven by poor first-order performance of patients with schizophrenia (Rouy et al. 2021). Similarly, in a recent meta-analysis, when young adults with the poorest recognition performance were compared to older adults with the best recognition performance, age-related episodic FOK disappeared (Devaluez et al. 2023).

This confound has been also strengthened by measures of metacognitive sensitivity in the memory domain. Most of these studies have used correlational methods (Marquié and Huet 2000, Souchay and Isingrini 2004, Dodson et al. 2007, Souchay et al. 2007, Thomas et al. 2011, Eakin and Hertzog 2012, Wong et al. 2012, Eakin et al. 2014, Sacher et al. 2015, Yeung 2024) such as gamma correlations (Goodman and Kruskal 1979). However, these measures are influenced by both memory performance and metacognitive bias (Masson and Rotello 2009, Fleming and Lau 2014, Vuorre and Metcalfe 2021). Thus, it is unclear whether the episodic FOK deficit observed in older adults is due to (i) a pure metamemory deficit, (ii) a memory deficit that would contaminate metamemory processes, and (iii) a memory deficit that would contaminate metacognitive sensitivity measure creating a spurious deficit.

To overcome this methodological limitation, metamemory sensitivity can be estimated using a model-based approach (Maniscalco and Lau 2012, 2014). This approach provides a measure of metacognitive efficiency (M_{ratio}), which is known to be more independent of first-order performance, compared to previous methods (Guggenmos 2021, Rahnev 2023). Using this framework, a few studies have focused on the evolution of RCJs in episodic and/or semantic memory across the lifespan with one finding no episodic metacognitive efficiency deficit in older adults (Palmer et al. 2014) and another a reduction in episodic metacognitive efficiency with aging (Meunier-Duperray et al. 2024). Importantly, a reduced episodic metacognition was only found when memory performance was also impaired with aging (Meunier-Duperray et al. 2024). This would suggest that episodic memory deficits would contaminate metamemory processes.

On the other hand, no study directly compared FOK judgments and RCJs in aging with the M_{ratio} framework. To our knowledge, only one study applied M_{ratio} to FOK judgments in young adults, showing reduced metacognitive efficiency compared to RCJs (Mazancieux et al. 2020a) as also typically found with other measures (Costermans et al. 1992). This study also revealed that whereas RCJs correlate across tasks, FOK judgments did not, which is also a pattern that has been unexplored in aging.

Thus, this study aimed at investigating the age effect on metamemory in a three novel way: (i) applying the M_{ratio} to FOK judgments in older adults to test whether previous episodic FOK deficit relates to the measure used, (ii) comparing FOK judgments and RCJs across the lifespan, and (3) replicating the deficit of episodic RCJs with aging (Meunier-Duperray et al. 2024) that remains unclear (Palmer et al. 2014) within the M_{ratio} framework. As preregistered, we hypothesized that metamemory efficiency for both FOK judgments and RCJs should be weaker in the episodic task for older adults compared to younger adults, whereas no difference in the semantic task should be observed. We also analyzed the effect of age on recall, recognition, and metacognitive bias, which may also be impacted (Castel et al. 2015, Culot et al. 2023, Kavé and Halamish 2015, Marquié and Huet 2000; nonpreregistered hypothesis).

Materials and methods

The preregistration is available on the Open Science Framework (OSF) website (https://osf.io/n3dma). The study received ethical guidance from the University Grenoble-Alpes (CERGA-Avis-2022-12). The experiment was based on a previous study that established differences in episodic–semantic FOK sensitivity and RCJs in younger adults (Mazancieux et al. 2020a).

Participants

Data were collected from 240 participants between 19 and 79 years. In the preregistration, we initially planned to only collect data for two groups: a younger adult group (age between 20 and 40 years old) and an older adult group (age between 60 and 80 years old). However, a recent version of the *HMeta-d* now allows the estimation of regression parameters on metacognitive efficiency

Table 1. Demographic details of the sample.

Age range (years)		18–35	35–59	60–79
n		119	40	81
Percentage of participants online		72.5%	100%	94.29%
Age (mean, SD)		23.00 (4.38)	44.80 (7.22)	68.70 (5.52)
Gender	Female	81%	70%	79%
	Male	17%	30%	21%
	Not given	2%	0%	0%
Years of education (mean, SD)		14.70 (3.18)	16.50 (3.12)	13.80 (3.31)
Cognitive score (mean, SD)		26.60 (2.03)	26.50 (2.16)	25.10 (2.36)

(Harrison et al. 2021). Therefore, we decided to also collect middle-aged participant data to treat age as a continuous variable and explore the evolution of metacognitive efficiency in adulthood.

Participants could take part in the experiment either online or in person at the University Grenoble-Alpes as a previous study conducted with large sample sizes has shown no effect of the type of testing on cognition and metacognition in both episodic and semantic tasks (Meunier-Duperray et al. 2024).

Table 1 summarizes the demographic data of the participants divided into three age groups (<35 years, between 35 and 59 years, and >59 years). Participants under the age of 35 years were mostly psychology students and were rewarded for their participation in the experiment with bonus points for their final exams. The three groups did not differ in terms of gender, $X^2(4) = 5.31$, $p = .26$, but they differed in terms of years of education, $F(2, 203) = 9.26$, $p < .01$. Middle-aged participants were more educated than the youngest participants ($d = -0.55$) and the oldest participants ($d = 0.81$). The number of years of education did not differ between the youngest and oldest participants ($d = 0.29$). All participants were native French speakers with normal or corrected vision and reported having no neurological or psychiatric disorders.

Before the experiment, all participants performed a quick cognitive test developed for this experiment to screen for (and measure) cognitive impairment for online testing. It was composed of exercises inspired by the Mini Mental State Examination (Folstein et al. 1975) and the Montreal Cognitive Assessment (Nasreddine et al. 2005). Five cognitive components were evaluated: orientation, memory, attention, language, and praxis. Participants scored an average of 25.79 (SD = 3.04) out of a total of 29 points. Participants whose scores were 2 SD below the overall mean (i.e., below 19.71/29) were considered to be cognitively impaired. Thus, one online younger adult and five older adults (three online participants) were removed from the analyses.

Materials

This study used the same stimuli as Mazancieux et al. (2020a). We created 80 sets of stimuli to be randomly allocated to the episodic (for word pairs) or semantic task (for word definitions). The cues of the semantic task were word definitions in French with the targets being the word associated with the definitions, and the distractor being a semantically associated word that did not correspond to the definition. Targets and distractors for the episodic task were the same target–distractor pairs as those for the semantic task. The cue and the target were semantically related. Thus, these 80 stimuli could be used as episodic or semantic material (Souchay et al. 2007) since each target could be related to a word of the episodic task or a definition of the semantic task. The stimuli were then divided into two counterbalanced sets of 40 stimuli each, randomly used as episodic or semantic stimuli for each participant. Participants who were presented with stimuli from the first list in the episodic task were presented with stimuli from the second list in the semantic task, and vice versa.

Procedure

All experimental sessions began with a brief description of the study and its goals, by the signature of the consent form, and the quick cognitive test. Then, the experiment was composed of two tasks implemented on PsychoPy (v2022.1.3) measuring metamemory in episodic and semantic memory independently. Participants performed the tasks on a computer with a black background screen and gave their responses with the keyboard and the mouse. Each task lasted ~20 min. Figure 1 summarizes the procedure of the two tasks.

In the episodic task, participants were presented with a 500-ms fixation cross followed by 40 semantically related pairs of words (a cue and a target) presented in random order on the left and on the right of the screen for 2500 ms each (e.g., meteorite and rock). Considering the possible memory difficulties of older adults, presentation time was adapted from the results of the previous article (presentation time of 1500 ms as in Mazancieux et al. 2020a). Participants were asked to memorize as many pairs as possible. Given that encouraging older adults to use semantic linking during encoding reduces the effect of age on memory (Guerrero Sastoque et al. 2019), participants were informed that the words in each pair were semantically linked and that this link could help them memorize the pairs. The position of the cue and the target were randomly assigned.

Following this encoding phase, a second phase displayed a 500-ms fixation cross followed by the same cues one by one in random order (e.g., rock). Participants had 15 s to write the target word using the keyboard. They could move to the next item after 5 s with a "validate" button. Whether the word was recalled or not, participants had to judge if they would be able to recognize the target between two words (FOK judgments). These judgments were made using a continuous scale from 50% ("I will answer randomly") to 100% ("I am sure I will recognize the correct answer") without limit of time. The following instruction appeared at the same time as the FOK scale: "In the next step, you will have to recognize the correct word from two propositions. How confident are you in recognizing the correct answer?"

In a last phase, participants were presented with all cues one by one in random order at the top of the screen. They had to choose between a target and a distractor which word was presented with the cue in the first phase (e.g., meteorite and comet by pressing the "F" or "H" key). Target and distractor positions were randomly assigned. After their responses, participants had to judge their confidence in their decision (RCJs). These judgments were made using the same continuous scale as for FOK judgments (from 50% to 100%). There was no time limit for both the decision and the RCJs. The RCJ scale appeared with the instruction: "How sure are you of your answer?" At the end of the task, participants could report if they used memory strategies (e.g., visualizing the words or creating a link between the two words). In the semantic task, participants were presented with 40 definitions one by one presented in random order (e.g., part of zoology that deals with birds). Participants had 15 s to write with the keyboard the word corresponding to the definition. They could move to the next item after 5 s with a "validate" button. Whether the word was recalled or not, participants had to judge if they would be able to recognize the

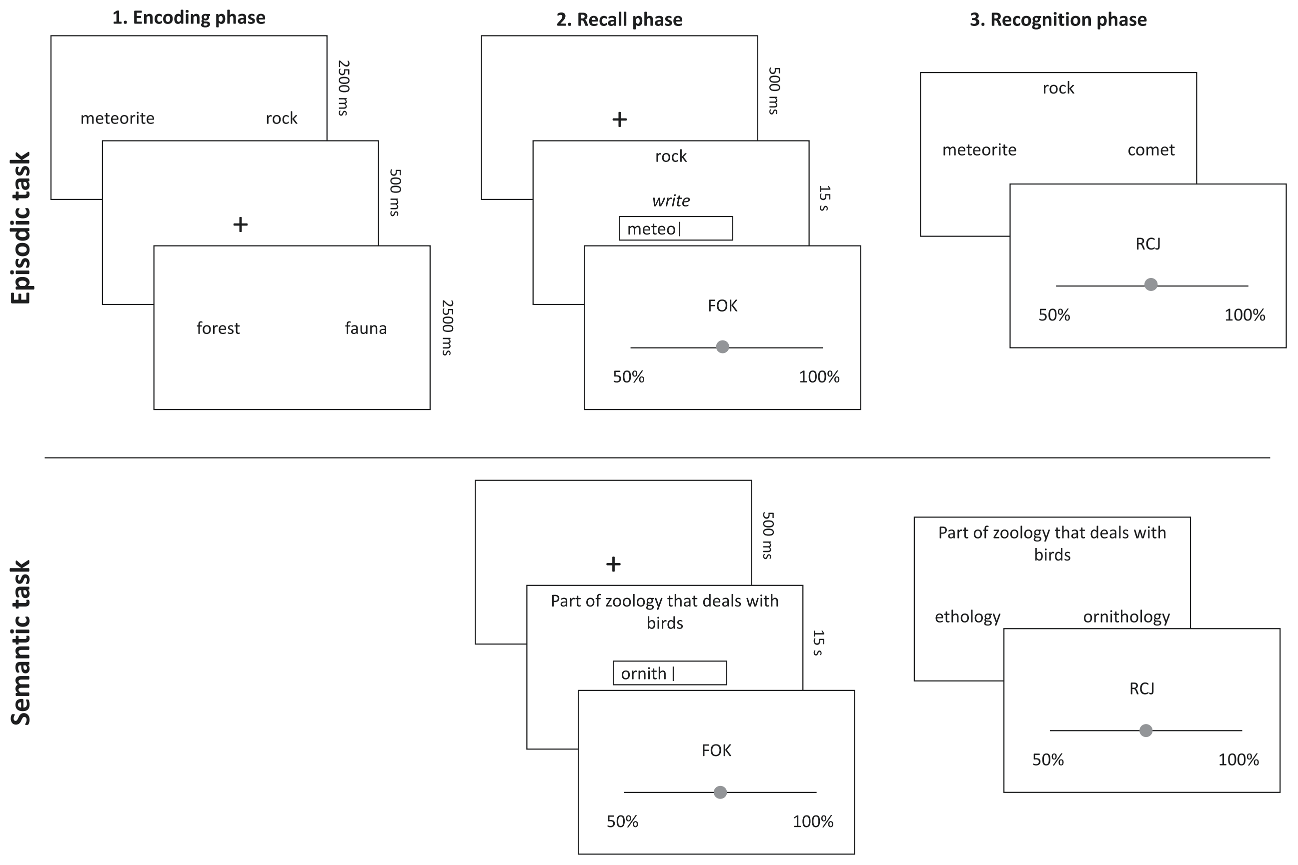

Figure 1. Procedure of the episodic and semantic memory tasks.

target between two words (FOK judgments using the same scale as for the episodic task). The following instruction was used for the FOK: "In the next step, you will have to recognize the correct word from two propositions. How confident are you in recognizing the correct answer?"

In a second phase, participants were presented with the same definitions one by one in a random order at the top of the screen. They had to recognize the right answer between a target and a distractor by pressing the "F" or "H" key (e.g., ethology and ornithology). The position of the correct response was randomly assigned. As for the episodic task, participants had to judge their confidence (RCJs) in their answer using the same continuous scale. There was no limit of time for both the decision and the RCJs. The RCJ scale was introduced with the instruction: "How sure are you of your answer?"

Before starting the task, written instructions were presented about how to respond for recall (i.e., write in a box), recognition (i.e., choose one answer from two), and on the judgment scale (i.e., move a red circle on a scale). Next, participants performed a test trial to ensure that they understood how to respond to each type of question. Finally, they were asked to provide an RCJ. At the end of the experiment, participants indicated their age, gender (male, female, or responded 'other'), and education level (years of schooling).

Data analyses and participant exclusions

Memory performance, metacognitive bias, and metacognitive efficiency

As preregistrered, memory performance and metacognitive bias were compared across the two age groups (20–40 and 60–80 years). As we decided to include middle-aged participants and to remain consistent with the Bayesian framework of *HMeta-d*, we also used Bayesian linear regression-type models for memory performance and metacognitive bias as in a previous study (Meunier-Duperray et al. 2024). As we preregistered frequentist analysis for these two dependent variables, these analyses have also been reported in the Supplementary material. These analyses were carried out using the Stan computational framework (http://mc-stan.org/) accessed with the brms package on R 4.2.2. The sampling procedure of the Stan framework is based on Markov chain Monte Carlo (MCMC) simulations. Models were run using four chains of 2000 iterations including 1000 warmup samples. The decision rule was based on the 95% highest density interval (HDI; Kruschke 2010), concluding to a high evidence for an effect when this HDI did not overlap with zero. For all measures, we tested the effects of age, task (episodic vs. semantic), judgment type (FOK judgments vs. RCJs), and interactions between age and the two other variables.

Memory performance was assessed with the proportion of correct recall and a sensitivity index (d') calculated on the recognition task. The following model was applied to estimate the effect of age on recall performance in each task (i.e., episodic and semantic), using default priors:

$$\text{recall} \sim \text{age} * \text{task}$$

For recognition performance, a sensitivity index (d') was calculated for each participant. To avoid infinite d' values from participants that recognized all items, the extreme proportions of hits were corrected (e.g., a hit proportion of 1 was converted to 0.975, see Macmillan and Creelman 2004). A sensitivity index of zero is the chance level, whereas a higher d' is associated with a better ability to discriminate between targets and distractors. We preregistered removing participants whose sensitivity index was at chance level (i.e., $d' \leq 0$). None of the participants had a random score for recognition. The following model was applied to estimate the effect of age on recognition performance in each task, using default priors:

$$d' \sim \text{age} * \text{task}$$

For the following analyses and as preregistered, we removed participants who used the same rating on the scale for all trials in each task independently by considering five breakpoints in the scale (i.e., [50–60[, [60–70[, [70–80[, [80–90[, and [90–100]). For the FOK judgments, 5 participants were removed from the semantic task and 12 for the episodic task. For the RCJs, six participants were removed from the semantic task and zero from the episodic task.

As memory performance varies across participants, metacognitive bias was estimated for each participant by the difference between their mean confidence and their percentage of correct responses in the recognition phase. Thus, a positive value suggests that participants overestimated their memory performance, whereas a negative value suggests that they underestimated their memory performance. The following model was applied to estimate the effect of age on recall performance in each task and for the two types of judgment (i.e., FOK judgments and RCJs), using default priors:

$$\text{bias} \sim \text{age} * \text{task} * \text{judgment}$$

As preregistered, metacognitive performance was estimated with a measure of the metacognitive efficiency based on signal detection theory (Maniscalco and Lau 2012, 2014). Using this framework, an ideal first-order performance value (*meta-d'*) is deduced from the observed confidence judgments of the participant (Maniscalco and Lau 2012, 2014). *Meta-d'* is then compared to the observed d' of the participant using a ratio between these two values (i.e., M_{ratio}). An M_{ratio} of 1 is considered as ideal metacognitive efficiency as observed and ideal values are the same. Here, we used a hierarchical Bayesian version of this approach (*HMeta-d*; Fleming 2017), which has the advantage of taking both the between- and within-subject variability into account. This approach allows a higher precision of group-level estimate particularly when the number of trials is low which is pertinent in the context of older adults. Bayesian computations are based on the hierarchical Bayesian model analysis program JAGS (http://mcmc-jags.sourceforge.net) and the sampling of posterior distributions of parameter estimates uses MCMC simulation. On the contrary to our preregistration that planned a comparison between the two age groups, we used an extended version adapted to estimate regression parameters (i.e., beta) (the *RHMeta-d*; Harrison et al. 2021). The continuous scales of metamemory judgments were converted to a five-point scale (i.e., [50-60[, [60–70[, [70–80[, [80–90[, and [90–100]). We tested the effect of age (using age z-scores) on the M_{ratio} in both memory tasks (i.e., episodic and semantic) and for both types of judgment (i.e., FOK judgments and RCJs) independently. We used the same parameters and priors as for the original *RHMeta-d* (Harrison et al. 2021).

Nonpreregistered exploratory analyses

We also explored the link between accuracy and confidence, drawing inspiration from the confidence–accuracy characteristic (CAC) analyses (Mickes 2015, Greene et al. 2024). These analyses consist in the estimation of the proportion of correct answers for each

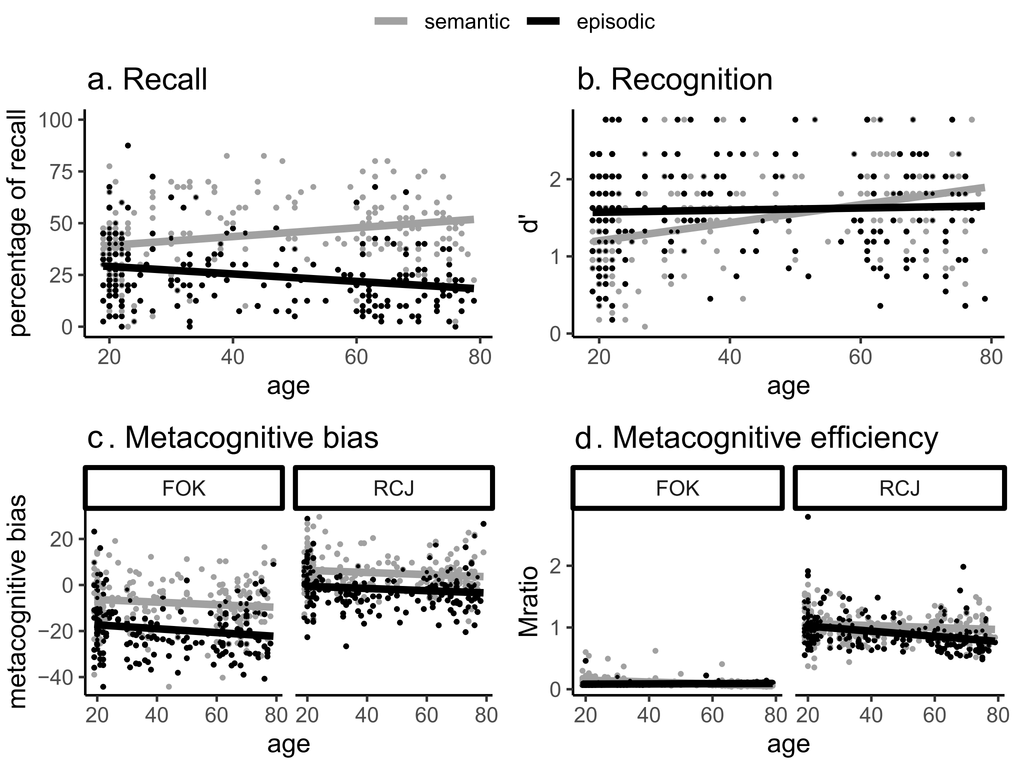

Figure 2. Distribution of (a) percent correct recall, (b) recognition performance (measured with d'), (c) metacognitive bias, and (d) metacognitive efficiency (M_{ratio}) and linear model regression line across age in episodic (black dots and lines) and semantic tasks (gray dots and lines) and for FOK judgments and RCJs.

confidence level. We performed these analyses on the proportion of correct recognition and recall, considering five confidence levels (i.e., [50–60], [60–70], [70–80], [80–90], and [90–100]). These analyses allowed us to characterize the response patterns and explain the observed M_{ratio}.

Results

Memory and metamemory performance were assessed across four measures: proportion of correct recall, recognition performance (with d'), metacognitive bias, and metacognitive efficiency (M_{ratio}). Figure 2 shows the distribution of data across age for each measure in the episodic and semantic tasks for FOK judgments and RCJs. Table 2 summarizes the mean and HDIs of the proportion of correct recall, recognition performance (d'), and metacognitive bias in the two memory domains (episodic and semantic) for the sample divided into three age groups and for all participants of the sample.

Memory performance

Memory performance was assessed with the proportion of correct recalls and a sensitivity index (d') calculated on the recognition task. For recall performance, there was no evidence for a difference between the episodic and semantic tasks ($\beta = -2.38$, HDI = [−8.83; 3.76]). Analyses showed evidence for a linear effect of age on recall performance ($\beta = 0.21$, HDI = [0.12; 0.31]), however with older adults recalling more items than younger adults. There was evidence for an interaction between age and task ($\beta = -0.39$, HDI = [−0.53; −0.26]): semantic recall increased with age, whereas episodic recall decreased with age.

For recognition, analyses showed evidence for better performance in the episodic task than in the semantic task ($\beta = 0.57$, HDI = [0.35; 0.80]). We also found an effect of age on recognition performance ($\beta = 0.012$, HDI = [0.008; 0.015]). As for recall performance, older adults performed better than younger adults. There was evidence for an interaction between age and task ($\beta = -0.010$, HDI = [−0.015; −0.006]) showing no effect of age on episodic recognition performance, but recognition performance for the semantic task increased with age. Note that preregistered frequentist analyses led to the same conclusions. Only the main effect of age on recall and the main effect of task on recognition were not significant. However, this does not change interpretation about the interactions that are the major results here.

Metacognitive bias

Results showed no effect of age on metacognitive bias ($\beta = -0.05$, HDI = [−0.11; 0.02]). In line with our hypothesis, metacognitive bias was weaker in the episodic task than in the semantic task ($\beta = -6.76$, HDI = [−11.09; −2.35]). There was evidence for an effect of judgment type ($\beta = -12.41$, HDI = [−17.00; −7.92]). In line with our hypothesis, participants were more underconfident with FOK judgments than with RCJs. Results showed no evidence for an interaction effect between age and task ($\beta = -0.00$, HDI = [−0.09; 0.08]) and between age and type of judgment ($\beta = -0.01$, HDI = [−0.10; 0.08]). There was no evidence for a three-way interaction between age, task, and type of judgment ($\beta = -0.02$, HDI = [−0.16; 0.11]).

Note that preregistered frequentist analyses led to the same conclusions. Only the main effect of the task (i.e., testing the difference between episodic and semantic tasks) on metacognitive bias was not significant. However, this does not change interpretation about the interactions that are the major results here.

Table 2. HDIs of proportion of correct recalls, recognition performance (d'), and metacognitive bias in the two memory domains (episodic and semantic) for younger participants (<35 years old), middle-aged participants (between 35 and 59 years old), older participants (over 59 years old), and overall sample participants.

	Proportion of correct recall		Recognition performance (d')		Metacognitive bias			
	Episodic	Semantic	Episodic	Semantic	Episodic FOK	Semantic FOK	Episodic RCJs	Semantic RCJs
<35 years old	0.28 [0; 0.63]	0.38 [0.08; 0.70]	1.56 [0.45; 2.77]	1.21 [0.18; 2.33]	−16.70 [−35.90; 9.79]	−5.80 [−25.40; 14.50]	−0.20 [−16.00; 18.00]	6.48 [−7.48; 26.70]
35–59 years old	0.26 [0.08; 0.48]	0.53 [0.33; 0.83]	1.76 [0.74; 2.77]	1.65 [0.85; 2.77]	−23.20 [−38.10; 1.99]	−8.88 [−44.20; 12.20]	−4.75 [−16.80; 6.61]	4.16 [−7.50; 20.40]
>59 years old	0.20 [0; 0.43]	0.48 [0.23; 0.80]	1.59 [0.55; 2.77]	1.73 [0.95; 2.77]	−20.40 [−39.00; 1.57]	−9.28 [−30.80; 10.50]	−1.88 [−20.50; 17.50]	4.35 [−5.81; 22.70]
All	0.25 [0; 0.58]	0.44 [0.15; 0.83]	1.60 [0.55; 2.77]	1.46 [0.64; 2.77]	−19.30 [−39.00; 2.24]	−7.70 [−28.60; 14.00]	−1.74 [−17.50; 15.10]	5.22 [−7.50; 23.20]

Note. For metacognitive bias, negative values indicate under-confidence, and positive values indicate over-confidence.

Table 3. Means and HDIs of metacognitive efficiency (M_{ratio}) in the two memory domains (episodic and semantic tasks) and for the two types of judgment (FOK judgments and RCJs).

	FOK judgments	RCJs
Episodic	0.07 [0.03; 0.12]	0.99 [0.92; 1.06]
Semantic	0.06 [0.03; 0.12]	0.86 [0.78; 0.94]

Note. The mean M_{ratio} and HDIs were estimated using the *HMeta-d* framework (Fleming 2017).

Metacognitive efficiency

Metacognition efficiency was estimated using M_{ratio}. The output from *RHMeta-d* allowed us to interpret the effect of task, type of judgment, and the effect of age in each memory task and type of judgment independently. The mean M_{ratio} and HDIs for each task and type of judgment are summarized in Table 3.

There was substantial evidence for an effect of the judgment type. The mean M_{ratio} was lower with FOK judgments than with RCJs in the both episodic and semantic tasks. There was no evidence of a task effect on metacognitive efficiency for either FOK judgments or RCJs. The mean M_{ratio} did not vary between episodic and semantic tasks.

There was no evidence of an age effect for episodic FOK ($\beta = -0.36$, HDI = [−0.92; 0.09]), semantic FOK ($\beta = 0.06$, HDI = [−0.55; 0.66]), and episodic RCJs ($\beta = -0.03$, HDI = [−0.09; 0.03]). There was substantial evidence for an effect of age on M_{ratio} in semantic RCJs ($\beta = -0.10$, HDI = [−0.18; −0.01]). Metacognitive efficiency decreased with age in the semantic task.

Exploratory analyses

As M_{ratio} for FOK was very low (although there was major evidence for a difference with zero), we checked that people were conducting this task appropriately by exploring the relationship between the distribution of metamemory judgments and the accuracy. To do so, we calculated the proportion of correct answers for each confidence level (as is common in the CAC approach; Mickes 2015). The difference between younger and older adults was not tested in these analyses since the foregoing analyses did not suggest any difference according to age.

These analyses showed that a higher proportion of correct recognitions was linked to higher RCJs in both episodic and semantic tasks, attesting that RCJs depend on response accuracy (Fig. 3a and b). However, this pattern was not observed in FOK judgments. The proportion of correct recognition was not linked to the FOK judgments in both episodic and semantic tasks (Fig. 3c and d), which is consistent with the low M_{ratio} values found for FOK judgments.

Relationship between FOK judgments and recall accuracy

The FOK task did not produce the typical pattern of metacognitive accuracy that has been demonstrated since Hart (1965), and raised the question of how our participants (many of them online) conducted the task. Previous work found correlations between FOK accuracy and recall performance, suggesting that recall can be used as a heuristic to answer FOK questions. We thus performed the same analyses as earlier with the proportion of correct recalls instead of recognition (Fig. 4). A clear pattern emerged showing that a higher proportion of correct recall was linked to higher FOK judgments in both episodic and semantic tasks, suggesting that

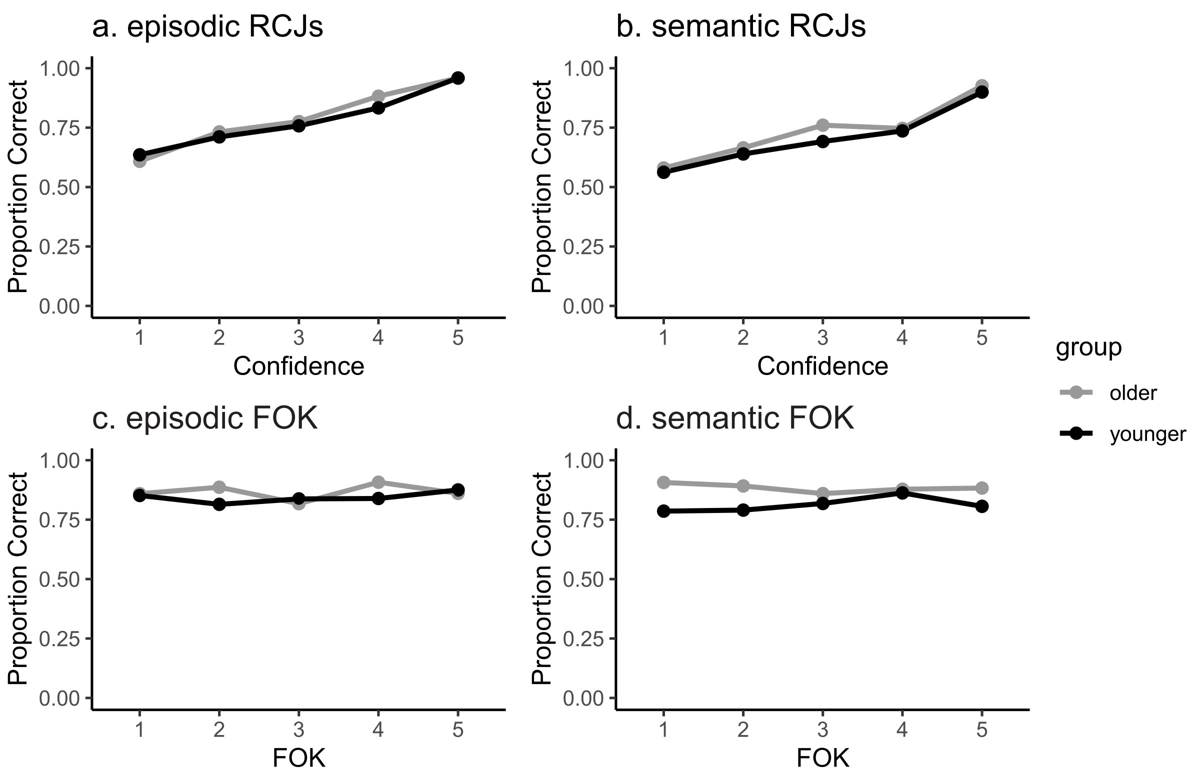

Figure 3. Relationship between the proportion of correct recognition and the confidence level for (a) episodic RCJs, (b) semantic RCJs, (c) episodic FOK judgments, and (d) semantic FOK judgments in younger adults (<60 years old) and older adults (>59 years old).

Note. The x-axis corresponds to the five breakpoints converted from the metacognitive judgments scale (i.e., [50–60[, [60–70[, [70–80[, [80–90[, and [90–100]). Thus, a value of 1 corresponds to low judgments (low FOK or low confidence) and a value of 5 corresponds to high judgments.

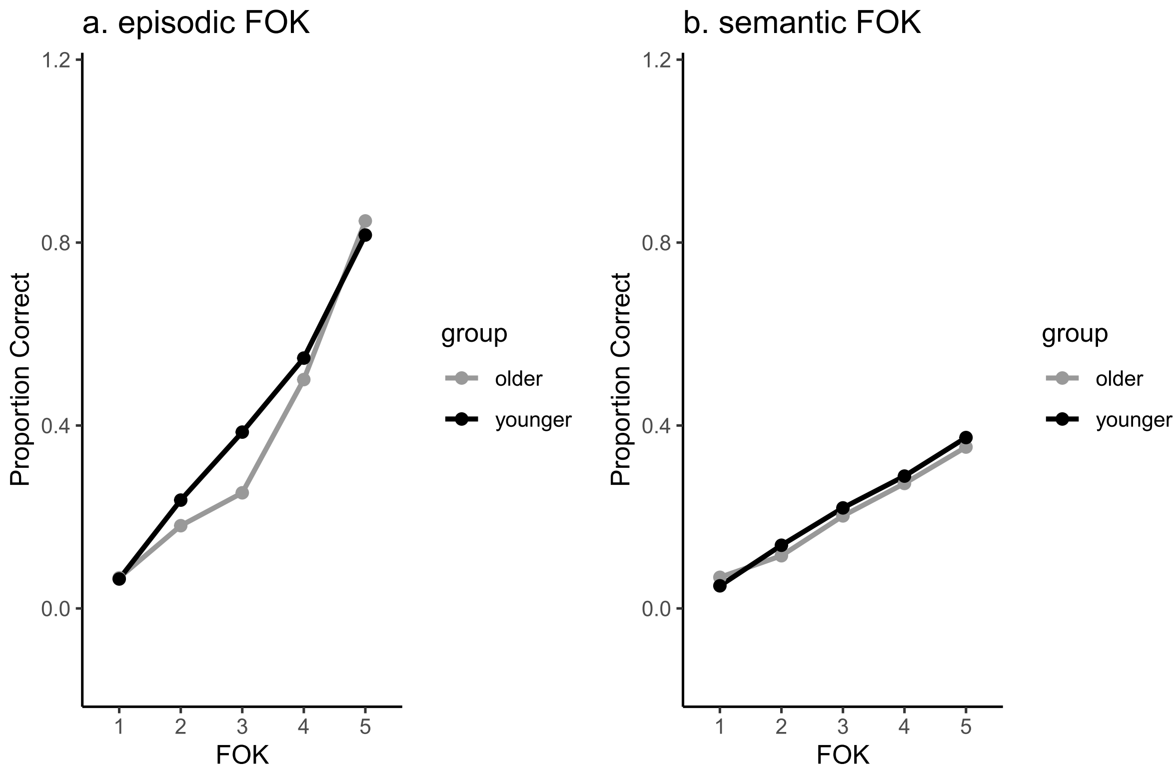

Figure 4. Relationship between the proportion of correct (prior) recall and the level of FOK in (a) episodic and (b) semantic tasks in younger adults (<60 years old) and older adults (>59 years old).

Note. The x-axis corresponds to the five breakpoints converted from the metacognitive judgments scale (i.e. [50–60[, [60–70[, [70–80[, [80–90[, and [90–100]). For the x-axis, a value of 1 corresponds to low judgments (low FOK or low confidence) and a value of 5 corresponds to high judgments.

FOK judgments were based on the previous recall accuracy rather than a prediction of future recognition.

Discussion

This study aimed at evaluating episodic and semantic metamemory in aging using FOK judgments and RCJs to provide a better understanding of dissociations observed within metamemory. Our main preregistered hypothesis was that metamemory efficiency for both FOK judgments and RCJs will be weaker in the episodic memory task for older adults compared to younger adults, whereas no difference was expected for metamemory efficiency between younger and older adults for the semantic memory task.

On the contrary to our hypothesis, our results did not show any effect of age on metacognitive efficiency in the episodic task. Although some studies have shown a negative effect of age on episodic metamemory using RCJs (Wong et al. 2012, Voskuilen et al. 2018, Meunier-Duperray et al. 2024), this effect is not always found (Dodson et al. 2007, Eakin et al. 2014, Palmer et al. 2014, Hertzog et al. 2021, Zakrzewski et al. 2021). A first explanation for these differences is that correct and incorrect responses must be treated separately when studying the metamemory abilities of older adults. Older adults are more likely to be highly confident about their memory errors than younger adults, while confidence judgments do not differ by age for correct responses (Chua et al. 2009, Soderstrom et al. 2012, Fandakova et al. 2013, Greene et al. 2022, 2024). However, the high confidence for errors was not shown in our exploratory analyses. Confidence judgments did not seem to differ between younger and older adults in terms of response accuracy (Fig. 3).

A second possibility is that metamemory abilities decrease in aging because of the decrease in episodic memory performance. In a previous study, we observed a decline in episodic metamemory when first-order performance was affected by aging (Meunier-Duperray et al. 2024). On the contrary here, we did not find an effect of aging on recognition memory nor an effect of aging on metacognitive efficiency. As previously mentioned, this can be explained by the fact that (i) an effective memory is necessary for accurate metamemory because both are partly based on the same processes and (ii) memory and metamemory are based on different processes but a memory deficit would contaminate measures of metacognitive sensitivity creating a spurious deficit in metacognitive sensitivity. Despite some unstable fits for very low first-order performance which we do not have in the current dataset, M_{ratio} has been shown to only have very weak dependence on task performance (Rahnev 2023). This suggests that the pattern observed in our previous study (Meunier-Duperray et al. 2024) and others (McWilliams et al. 2023) is likely due to shared processes between memory and metamemory. If M_{ratio} has appropriate psychometric measures, this should be taken into account in the measure. Thus, the SDT model assumes that both first- and second-order responses are based on the same underlying evidence (Maniscalco and Lau 2012). Considering that M_{ratio} correctly controls first-order performance, the decrease in metacognitive efficiency observed by Meunier-Duperray et al. (2024) in a task where first-order performance also decreases with age can be explained by the fact that memory decline has a greater impact on metacognition than recognition memory itself. This is supported by the lognormal meta-noise model of metacognition (Shekhar and Rahnev 2021) in which metacognitive noise is signal dependent increasing for extreme evidence values.

In line with previous literature (Mitchell 1989, Nyberg et al. 1996, Cosgrove et al. 2021), the proportion of correct recall increased with age in the semantic task, whereas it decreased with age in the episodic task. In contrast, recognition performance increased with age in the semantic task, while it did not vary with age in the episodic task. This dissimilar pattern between recall and recognition for the episodic memory task has been previously reported (Craik and McDowd 1987) and can be explained by processes involved in recognition memory. Recall is more resource demanding than recognition and decreasing processing resources occurs with aging. Moreover, recall relies primarily on the recollection process (Mandler 1980), which is underused by older adults (Jacoby 1999, Daselaar et al. 2006). During recognition, older adults may rely exclusively on the familiarity process, which is preserved in aging (Mandler 1980, Jacoby 1999, Daselaar et al. 2006). As it has been shown that the quality and quantity of recollection cues (Thomas et al. 2011, Souchay and Isingrini 2012) improve metacognitive sensitivity, this can also explain that when recollection is impaired as in aging, first-order performance may be less impacted than second-order performance creating the episodic RCJ pattern observed in Meunier-Duperray et al. (2024).

We suggest that this pattern was not replicated here because of the nature of the episodic memory task. Participants were informed before the encoding stage that the paired words in the episodic task were semantically related and that this could help memorize the pairs. Older adults have difficulty spontaneously using environmental support during memory tasks, but they are able to adapt their strategies in order to improve memory performance as soon as they are encouraged to use semantic processing during encoding (Guerrero Sastoque et al. 2019). This instruction was intended to facilitate encoding and reduce the age difference in episodic performance in order to reduce variations in first-order performance. We propose here that this may have been also used as a metacognitive cue by older adults.

Surprisingly, we found an effect of age on metacognitive efficiency for semantic RCJs, which is not typically found in semantic tasks (Lachman et al. 1979, Marquié and Huet 2000, Dodson et al. 2007, Dahl et al. 2009, Eakin et al. 2014, Kavé and Halamish 2015). This is also in contradiction with Meunier-Duperray et al. (2024) where no metacognitive efficiency deficit was found. One possible explanation for the differences between these two studies is the nature of the semantic tasks. Whereas we previously used general knowledge questions, the present study used definitions. The cues used to rate confidence for lexical material may differ according to age-related changes in the semantic network related to language production (Martin et al. 2022). This remains to be replicated in a context where different semantic tasks are used. At a more theoretical level, these results highlighted a clear dissociation between semantic memory and metacognitive efficiency: memory performance increased with age, while metacognition decreased. This suggests that some metacognitive processes are completely independent from first-order processes leading to metacognitive inefficiency in aging. Moreover, these processes would be domain specific as we found no age effect on episodic RCJs (see Mazancieux et al. (2023) for a review about domain-general metacognition).

Turning to FOK results, we found no evidence of an age effect on metacognitive efficiency using FOK judgments in both episodic and semantic tasks. Although our estimates of metacognitive efficiency were very low, the mean M_{ratio} was above chance ($\mu M_{ratio} = 0.07$ and $\mu M_{ratio} = 0.06$ in the episodic and semantic FOK tasks, respectively), demonstrating weak but measurable metacognitive abilities in our participants. This finding contradicts previous studies, which have often shown an age effect in episodic FOK tasks, but not in semantic FOK tasks (see Devaluez et al. (2023) for a review and meta-analysis).

Exploratory analyses showed that FOK judgments were based on recall accuracy rather than a prediction of future recognition. This result is similar to Mazancieux et al. (2020a) where FOK ratings (both episodic and semantic) were better explained by the accuracy of the recall task compared to the accuracy of the recognition task. Recall was also positively correlated with FOK metacognitive efficiency. However, in this study, recognition accuracy was significantly predicted by FOK judgments and M_{ratio} were largely higher than in the current study (mean $M_{ratio} = 0.65$ in the episodic FOK task and $M_{ratio} = 0.55$ in the semantic FOK task). The main difference between the two studies is that here 89% of the participants performed the task online. To provide a comparable context between online and in-person testing, participants who performed the task in person were left alone in a room during the tasks. The experimenter was present only for the short cognitive test and during the practice trial of the experimental task, and then left the room. A related possibility is that written instructions given to the participant were not clear enough. Instructions were as follows: "In the next step, you will have to recognize the correct word from two propositions. How confident are you in recognizing the correct answer?" The term "confident" was perhaps too close to the RCJ instructions, which made the participant interpret FOK as the RCJ for recall.

We expected an age-related decline in the metacognitive efficiency of the episodic FOK task, not least because it involves the recollection process (e.g., Hicks et al. 2002), in particular the retrieval of the encoding context (Hertzog et al. 2014, Isingrini et al. 2016). No effect of aging was found here which might be due to the fact that RCJs and FOK judgments are made within the same task. In line with our results, Eakin et al. (2014) found no effect of age on both episodic and semantic FOK sensitivity (i.e., gamma scores) using an experimental design that included both FOK judgments and RCJs in the same task. In sum, performing RCJs may have helped older adults to perform the FOK task. The effect of retrospective judgment magnitude on prospective judgment has also been identified in visual perception (Fleming et al. 2016). Moreover, the effects of FOK judgments on RCJs have also been observed (Pournaghdali 2022). Another very likely possibility is that M_{ratio} related to FOK judgments was too low even for young adults to identify any decrease with age. Note that in Eakin et al. (2014), moderate levels of FOK accuracy in both episodic (gamma mean = 0.39; SD = 0.36) and semantic tasks (gamma mean = 0.51; SD = 0.49) were found.

Regarding metacognitive bias, our results are mixed. Our Bayesian analyses showed that substantial evidence is in favor of a negative effect of aging regardless of tasks and types of judgments ($\beta = -0.05$, HDI = [−0.11; 0.02]). However, HDI slightly overlaps with zero, which makes these results inconclusive. A domain-general decrease in the magnitude of RCJs with aging across memory and perception has been previously reported (McWilliams et al. 2023). Our data suggest that this might also be the case here. Other studies also showed that older adults give lower FOK judgments than younger adults (Marquié and Huet 2000, Castel et al. 2015) but are well calibrated for RCJs (Kavé and Halamish 2015). It is possible that different profiles of normal aging can be at play leading to potential variability within and across samples. For instance, it has been shown that higher levels of cognitive reserve allow individuals to use effective memory strategies and control (Frankenmolen et al. 2018, Angel et al. 2022).

To conclude, we showed that metacognitive efficiency related to semantic RCJs decreased in aging, while metacognitive efficiency related to episodic RCJs was stable, contradicting previous results (showing a decline of episodic metamemory and preservation of semantic metamemory). For our episodic task, we do not find an age effect on metacognitive efficiency, which can be rationalized according to differences between studies where there are and are not age differences in memory performance. For our semantic task, however, we showed that metacognitive abilities can be impaired even when first-order performance is well preserved, highlighting the importance of having a measure of metacognitive abilities that is independent of first-order performance. These results point to the complexity of the cues used and processes involved in metacognition for memory tasks, and seem to contradict the notion of domain generality in metacognition. Surprisingly, the metacognitive efficiency for FOK judgments was very low, highlighting the poor metacognitive performance of participants, regardless of age and memory domain. We found that participants based their FOK judgments on previous recall accuracy rather than on a prediction of future recognition. Our study may also point to differences between online and in-person study results for tasks that need more instructions and an experimenter for participants to ask questions. Participants reinterpreted the FOK task, making our FOK results less interesting for understanding the aging pattern, but offering novel insights into how participants complete metacognitive judgments and interpret our experimental tasks. Finally, future work may look at other factors that may contribute to decrease in both metacognitive bias and efficiency such as cognitive reserve.

Author contributions

Lucile Meunier-Duperray (Software, Formal analysis, Investigation, Data curation, Writing—original draft, Visualization), Audrey Mazancieux (Conceptualization, Writing—review & editing), Céline Souchay (Conceptualization, Writing—review & editing, Supervision), Christine Bastin (Conceptualization, Writing—review & editing), Chris J.A. Moulin (Conceptualization, Writing—review & editing, Supervision, Funding acquisition), and Lucie Angel (Conceptualization, Writing—review & editing, Supervision)

Supplementary data

Supplementary data is available at *Neuroscience of Consciousness* online.

Conflict of interest

None declared.

Funding

This work was financed by an *Agence National de Recherche* award to C.J.A.M, L.A., C.B., and C.S., AGEFOK: Understanding metamemory in healthy aging, ANR-21-CE28-0002-01.

Data availability

Raw data, data scripts, and other materials (https://osf.io/ymg96/), as well as the preregistration (https://osf.io/n3dma), are available on OSF.

References

Allen-Burge R, Storandt M. Age equivalence in feeling-of-knowing experiences. *J Gerontol B Psychol Sci Soc Sci* 2000;**55**:P214–23.

Angel L, Guerrero-Sastoque L, Bernardo M *et al.* Metamemory mediates the protective effect of cognitive reserve on episodic memory during aging. *Acta Psychol* 2022;**228**:103627.

Bastin C, Giacomelli F, Miévis F *et al.* Anosognosia in mild cognitive impairment: lack of awareness of memory difficulties characterizes prodromal Alzheimer's disease. *Front Psychiatry* 2021;**12**:631518.

Castel AD, Middlebrooks CD, and McGillivray S. Monitoring memory in old age: impaired, spared, and aware. In: Dunlosky J and Tauber S (Uma) K (eds.), *The Oxford Handbook of Metamemory*, Vol. **1**. UK: Oxford University Press, 2015, 519–35.

Chua EF, Schacter DL, Sperling RA. Neural basis for recognition confidence in younger and older adults. *Psychol Aging* 2009;**24**:139–53.

Cosgrove AL, Kenett YN, Beaty RE *et al.* Quantifying flexibility in thought: the resiliency of semantic networks differs across the lifespan. *Cognition* 2021;**211**:104631.

Costermans J, Lories G, Ansay C. Confidence level and feeling of knowing in question answering: the weight of inferential processes. *J Exp Psychol* 1992;**18**:142–50.

Craik FIM, McDowd JM. Age differences in recall and recognition. *J Exp Psychol* 1987;**13**:474–9.

Culot C, Lauwers T, Fantini-Hauwel C *et al.* Contributions of age and clinical depression to metacognitive performance. *Consciou Cognit* 2023;**107**:103458.

Dahl M, Allwood CM, Hagberg B. The realism in older people's confidence judgments of answers to general knowledge questions. *Psychol Aging* 2009;**24**:234–8.

Daselaar SM, Fleck MS, Dobbins IG *et al.* Effects of healthy aging on hippocampal and rhinal memory functions: an event-related fMRI study. *Cereb Cortex* 2006;**16**:1771–82.

Devaluez M, Mazancieux A, Souchay C. Episodic and semantic feeling-of-knowing in aging: a systematic review and meta-analysis. *Sci Rep* 2023;**13**:1.

Dodson CS, Bawa S, Krueger LE. Aging, metamemory, and high-confidence errors: a misrecollection account. *Psychol Aging* 2007;**22**:122–33.

Eakin DK, Hertzog C. Age invariance in feeling of knowing during implicit interference effects. *J Gerontol Ser B* 2012;**67**:555–62.

Eakin DK, Hertzog C, Harris W. Age invariance in semantic and episodic metamemory: both younger and older adults provide accurate feeling-of-knowing for names of faces. *Aging Neuropsychol Cognit* 2014;**21**:27–51.

Elman JA, Klostermann EC, Marian DE *et al.* Neural correlates of metacognitive monitoring during episodic and semantic retrieval. *Cognit Affective Behav Neurosci* 2012;**12**:599–609.

Fandakova Y, Shing YL, Lindenberger U. High-confidence memory errors in old age: the roles of monitoring and binding processes. *Memory* 2013;**21**:732–50.

Flavell JH. First discussant's comments: what is memory development the development of? *Human Dev* 1971;**14**:272–8.

Fleming SM. HMeta-d: hierarchical Bayesian estimation of metacognitive efficiency from confidence ratings. *Neurosci Consciou* 2017;**2017**:nix007.

Fleming SM, Lau HC. How to measure metacognition. *Front Human Neurosci* 2014;**8**:443.

Fleming SM, Massoni S, Gajdos T *et al.* Metacognition about the past and future: quantifying common and distinct influences on prospective and retrospective judgments of self-performance. *Neurosci Consciou* 2016;**2016**:niw018.

Folstein S, Folstein SE, McHugh PR. "Mini-mental state." A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;**12**:189–98.

Frankenmolen NL, Fasotti L, Kessels RPC *et al*. The influence of cognitive reserve and age on the use of memory strategies. *Expl Aging Res* 2018;**44**:117–34.

Goodman LA, and Kruskal WH. *Measures of Association for Cross Classifications*. New York: Springer New York, 1979.

Greene NR, Chism S, Naveh-Benjamin M. Levels of specificity in episodic memory: insights from response accuracy and subjective confidence ratings in older adults and in younger adults under full or divided attention. *J Exp Psychol Gen* 2022;**151**:804–19.

Greene NR, Forsberg A, Guitard D *et al*. A lifespan study of the confidence-accuracy relation in working memory and episodic long-term memory. *J Exp Psychol Gen* 2024;**153**:1336–60.

Guerrero Sastoque LF, Bouazzaoui B, Burger L *et al*. Optimizing memory strategy use in young and older adults: the role of metamemory and internal strategy use. *Acta Psychol* 2019;**192**:73–86.

Guggenmos M. Measuring metacognitive performance: type 1 performance dependence and test-retest reliability. *Neurosci Consciou* 2021;**2021**:niab040.

Harrison OK, Garfinkel SN, Marlow L *et al*. The Filter Detection Task for measurement of breathing-related interoception and metacognition. *Biol Psychol* 2021;**165**:108185.

Hart JT. Memory and the feeling-of-knowing experience. *J Educ Psychol* 1965;**56**:208–16.

Hertzog C, Curley T, Dunlosky J. Are age differences in recognition-based retrieval monitoring an epiphenomenon of age differences in memory? *Psychol Aging* 2021;**36**:186–99.

Hertzog C, Fulton EK, Sinclair SM *et al*. Recalled aspects of original encoding strategies influence episodic feelings of knowing. *Memory Cognit* 2014;**42**:126–40.

Hicks JL, Marsh RL. On predicting the future states of awareness for recognition of unrecallable items. *Memory Cognit* 2002;**30**:60–6.

Hicks JL, Marsh RL, Ritschel L. The role of recollection and partial information in source monitoring. *J Exp Psychol* 2002;**28**:503–8.

Irak M, Soylu C, Yavuz M. Comparing event-related potentials of retrospective and prospective metacognitive judgments during episodic and semantic memory. *Sci Rep* 2023;**13**:1949.

Isingrini M, Sacher M, Perrotin A *et al*. Episodic feeling-of-knowing relies on noncriterial recollection and familiarity: evidence using an online remember-know procedure. *Consciou Cognit* 2016;**41**:31–40.

Jacoby LL. Ironic effects of repetition: measuring age-related differences in memory. *J Exp Psychol* 1999;**25**:3–22.

Jarjat G, Ward G, Hot P *et al*. Distinguishing the impact of age on semantic and nonsemantic associations in episodic memory. *J Gerontol B* 2021;**76**:722–31.

Kavé G, Halamish V. Doubly blessed: older adults know more vocabulary and know better what they know. *Psychol Aging* 2015;**30**:68–73.

Kikyo H, Ohki K, Miyashita Y. Neural correlates for feeling-of-knowing: an fMRI parametric analysis. *Neuron* 2002;**36**:177–86.

Kruschke JK. Bayesian data analysis. *Wiley Interdiscip Rev Cogn Sci* 2010;**1**:658–76.

Lachman JL, Lachman R, Thronesbery C. Metamemory through the adult life span. *Dev Psychol* 1979;**15**:543–51.

Lund AE, Correa C, Fardo F *et al*. Domain generality in metacognitive ability: a confirmatory study across visual perception, memory, and general knowledge. 2023.

Macmillan NA, and Creelman CD. *Detection Theory: A User's Guide*. New Jersey: Psychology Press, 2004.

Mandler G. Recognizing: the judgment of previous occurrence. *Psychol Rev* 1980;**87**:252–71.

Maniscalco B, Lau H. A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciou Cognit* 2012;**21**:422–30.

Maniscalco B, and Lau H. Signal detection theory analysis of type 1 and type 2 data: Meta-d′, response-specific Meta-d′, and the unequal variance SDT model. In: Fleming SM and Frith CD (eds.), *The Cognitive Neuroscience of Metacognition*. Berlin: Springer, 2014, 25–66.

Marquié JC, Huet N. Age differences in feeling-of-knowing and confidence judgments as a function of knowledge domain. *Psychol Aging* 2000;**15**:451–61.

Martin S, Saur D, Hartwigsen G. Age-dependent contribution of domain-general networks to semantic cognition. *Cereb Cortex* 2022;**32**:870–90.

Masson MEJ, Rotello CM. Sources of bias in the Goodman–Kruskal gamma coefficient measure of association: implications for studies of metacognitive processes. *J Exp Psychol* 2009;**35**:509–27.

Mazancieux A, Dinze C, Souchay C *et al*. Metacognitive domain specificity in feeling-of-knowing but not retrospective confidence. *Neurosci Consciou* 2020a;**2020**:1–11.

Mazancieux A, Fleming SM, Souchay C *et al*. Is there a G factor for metacognition? Correlations in retrospective metacognitive sensitivity across tasks. *J Exp Psychol Gen* 2020b;**149**: 1788–99.

Mazancieux A, Pereira M, Faivre N *et al*. Towards a common conceptual space for metacognition in perception and memory. *Nat Rev Psychol* 2023;**2**:751–766.

McWilliams A, Bibby H, Steinbeis N *et al*. Age-related decreases in global metacognition are independent of local metacognition and task performance. *Cognition* 2023;**235**:105389.

Meunier-Duperray L, Mazancieux A, Souchay C *et al* (2024). *Does Age Affect Metacognition? A Cross-domain Investigation Using a Hierarchical Bayesian Framework*..

Mickes L. Receiver operating characteristic analysis and confidence–accuracy characteristic analysis in investigations of system variables and estimator variables that affect eyewitness memory. *J Appl Res Memory Cognit* 2015;**4**:93–102.

Mitchell DB. How many memory systems? Evidence from aging. *J Exp Psychol* 1989;**15**:31–49.

Morson SM, Moulin CJA, Souchay C. Selective deficits in episodic feeling of knowing in ageing: a novel use of the general knowledge task. *Acta Psychol* 2015;**157**:85–92.

Nasreddine ZS, Phillips NA, Bédirian V *et al*. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc* 2005;**53**:695–9.

Nelson TO, and Narens L. Metamemory: a theoretical framework and new findings. In: *Psychology of Learning and Motivation*. Vol. **26**, Cambrdige: Academic Press, 1990, 125–173.

Nyberg L, Bäckman L, Erngrund K *et al*. Age differences in episodic memory, semantic memory, and priming: relationships to demographic, intellectual, and biological factors. *J Gerontol B* 1996;**51B**:234–P240.

Palmer EC, David AS, Fleming SM. Effects of age on metacognitive efficiency. *Consciou Cognit* 2014;**28**:151–60.

Pannu JK, Kaszniak AW. Metamemory experiments in neurological populations: a review. *Neuropsychol Rev* 2005;**15**:105–30.

Perfect TJ, Stollery B. Memory and metamemory performance in older adults: one deficit or two? *Quarterly J Exp Psychol Sec A* 1993;**46**:119–35.

Perrotin A, Isingrini M, Souchay C *et al*. Episodic feeling-of-knowing accuracy and cued recall in the elderly: evidence for double dissociation involving executive functioning and processing speed. *Acta Psychol* 2006;**122**:58–73.

Pournaghdali A. Signal Detection Analyses of the Relation of Prospective and Retrospective Metacognitive Judgments. Florida International University, 2022.

Rahnev D. *Measuring Metacognition: A Comprehensive Assessment of Current Methods*. PsyArXiv 2023.

Reggev N, Zuckerman M, Maril A. Are all judgments created equal?: An fMRI study of semantic and episodic metamemory predictions. *Neuropsychologia* 2011;**49**:1332–42.

Rouy M, Saliou P, Nalborczyk L *et al*. Systematic review and meta-analysis of metacognitive abilities in individuals with schizophrenia spectrum disorders. *Neurosci Biobehav Rev* 2021;**126**:329–37.

Sacher M, Isingrini M, Taconnat L. Effects of aging and divided attention on episodic feeling-of-knowing accuracy. *Acta Psychol* 2013;**144**:258–63.

Sacher M, Landré L, Taconnat L. Age-related differences in episodic feeling-of-knowing arise from differences in memory performance. *Memory* 2015;**23**:119–26.

Shekhar M, Rahnev D. The nature of metacognitive inefficiency in perceptual decision making. *Psychol Rev* 2021;**128**: 45–70.

Soderstrom NC, McCabe DP, Rhodes MG. Older adults predict more recollective experiences than younger adults. *Psychol Aging* 2012;**27**:1082–8.

Souchay C, Isingrini M. Age-related differences in the relation between monitoring and control of learning. *Expl Aging Res* 2004;**30**:179–93.

Souchay C, Isingrini M. Are feeling-of-knowing and judgment-of-learning different? Evidence from older adults. *Acta Psychol* 2012;**139**:458–64.

Souchay C, Moulin CJA, Clarys D *et al*. Diminished episodic memory awareness in older adults: evidence from feeling-of-knowing and recollection. *Consciou Cognit* 2007;**16**:769–84.

Thomas AK, Bulevich JB, Dubois SJ. Context affects feeling-of-knowing accuracy in younger and older adults. *J Exp Psychol* 2011;**37**:96–108.

Voskuilen C, Ratcliff R, McKoon G. Aging and confidence judgments in item recognition. *J Exp Psychol* 2018;**44**:1–23.

Vuorre M, and Metcalfe J. Measures of relative metacognitive accuracy are confounded with task performance in tasks that permit guessing. *Metacogn Learn* 2021;17:269-291.

Wong JT, Cramer SJ, Gallo DA. Age-related reduction of the confidence–accuracy relationship in episodic memory: effects of recollection quality and retrieval monitoring. *Psychol Aging* 2012;**27**:1053–65.

Yeung MK. Metamemory and executive function mediate the age-related decline in memory. *J Int Neuropsychol Soc* 2024:30:479–488.

Zakrzewski AC, Sanders EC, Berry JM. Evidence for age-equivalent and task-dissociative metacognition in the memory domain. *Front Psychol* 2021;**12**:630143.

Neuroscience of Consciousness, 2025, 2025(1), niaf003 , DOI: https://doi.org/10.1093/nc/niaf003, Research Article
Received 27 August 2024; Revised 24 December 2024; Accepted 10 February 2025