Metacognitive domain specificity in feeling-of-knowing but not retrospective confidence

Audrey Mazancieux *, Claire Dinze, Céline Souchay and Chris J.A. Moulin

Laboratoire de Psychologie et NeuroCognition (LPNC) CNRS 5105, 1251 Avenue Centrale, St Martin d’Hères, Université Grenoble Alpes, 38040 Grenoble, France

*Corresponding author. Laboratoire de Psychologie et NeuroCognition (LPNC) CNRS 5105, 1251 Avenue Centrale, St Martin d’Hères, Université Grenoble Alpes, 38040 Grenoble, France. Tel: +33 4 76 82 57 10; Email:audrey.mazancieux@gmail.com

Abstract

Previous research has converged on the idea that metacognitive evaluations of memory dissociate between semantic and episodic memory tasks, even if the type of metacognitive judgement is held constant. This often observed difference has been the basis of much theoretical reasoning about the types of cues available when making metacognitive judgements of memory and how metacognition is altered in memory pathologies. Here, we sought to revisit the difference between episodic and semantic feeling-of-knowing (FOK) judgements in the light of recent research which has supported a domain general account of metacognition. One hundred participants performed classical episodic and semantic memory tasks with FOK judgements and confidence judgements. Using the meta-d0 framework, we applied a hierarchical Bayesian model to estimate metacognitive efficiency (meta-d0/d0) between the episodic memory task and the semantic memory task for confidence judgements; however, no evidence was found for a cross-task correlation for FOK judgements. This supports the view that FOK judgements are based on different cues in semantic and episodic memory, whereas confidence judgements are domain general.

Keywords: metacognition; feeling-of-knowing; confidence; episodic memory; semantic memory

Introduction

When people fail to retrieve information from memory, they may have a feeling that they nonetheless know this information. Such feelings-of-knowings (FOKs) are a self-evaluation of cognitive abilities and can be thought of as a metacognitive experience, akin to the tip-of-the-tongue state (see Moulin and Souchay 2014). In experiments, FOKs refer to the predictions of future stimulus recognition when this stimulus has not been recalled. They are a special kind of judgement in that they are cued by a retrieval attempt but pertain to future performance, and in this sense, they are prospective judgements. In the episodic FOK (eFOK) paradigm (e.g. Schacter 1983; Souchay et al. 2000), participants learn cue-target paired-words. In a subsequent recall phase, a cue word is presented and the participants are asked to recall the target word. In cases where participants are not able to remember the target, they judge whether they will be able to recognize the target amongst a set of distracter words (the FOK prediction). Finally, participants perform the recognition task. In the semantic FOK (sFOK) paradigm (e.g. Hart 1965; Nelson et al. 1990) instead of learning paired-words, word definitions or general knowledge questions are presented to participants (e.g. “A movement, passage, or composition marked to be performed in slow time”) and they are asked to recall the answer or word referring to this definition (adagio). As in the episodic task, they make an FOK if they are not able to find the correct word, and later perform a recognition task. In the current experiment, we address the question of whether sFOK and eFOK share similar basis or resources.

The literature shows striking dissociations between the accuracy of judgements in sFOK and eFOK and especially in what is named “metacognitive sensitivity” (i.e. the ability to...
discriminate between correct and incorrect responses). In several neuropsychological populations, a profile of impaired eFOK despite preserved sFOK has been reported (e.g. Alzheimer’s disease, Souchay 2007; schizophasia, Bacon et al. 2001; Souchay et al. 2006; patients with frontal lobe lesions, Schnyer et al. 2004). The same dissociation is observed in older adults (e.g. Souchay et al. 2007). This profile is proposed to occur because sFOK and eFOK are based on different retrieval processes. Hicks et al. (2002) suggest that eFOK rely on autonoetic consciousness in order to retrieve partial information from the study phase on which to base the FOK prediction such as being able to recall what you were thinking when you first saw the pair. Such a process is exclusively involved in episodic memory. In comparison, sFOK judgements do not rely on autonoetic consciousness but can be inferred from the retrieval of lexical or semantic information associated to the target question such as when trying to answer the question: “word obtained by transposing the letters of another word,” one can know that the word ends with “gram.” Finally, in terms of neural regions, sFOK and eFOK seem to be based on both common and distinct regions (Reggev et al. 2011; Elman et al. 2012).

This distinction is a critical one in theories of metacognition and memory, helping identify the cues used to make metacognitive evaluations of retrieval, and pointing to a domain specificity of FOK accuracy in human memory. Despite being such a robust finding in the neuropsychological and aging literatures, to our knowledge no research has examined the episodic–semantic distinction in correlational tasks in healthy participants. This is of interest for two reasons which we develop below. First, it seems that metacognitive failure for episodic materials only tends to occur in special populations with an episodic memory deficit (the Memory Constraint Hypothesis; Hertzog et al., 2010) limiting the generalizability of this theory. Moreover, a number of methodological issues raise questions about the nature of the episodic–semantic dissociation (for a recent review, see Renoult et al. 2019). Secondly, research in other cognitive domains has found evidence of domain generalizable processes in metacognition (e.g. McCurdy et al. 2013; Ais et al. 2016; Samaha and Postle, 2017; Lee et al. 2018; Mazancieux et al. 2018), contrasting with the observed neuropsychological dissociation between semantic and episodic memory. We introduce these two areas in turn.

Firstly, several methodological issues cloud the results of studies comparing sFOK and eFOK. As shown in Table 1, few studies have directly compared these two types of FOK and most of the dissociations related above are driven by distinct studies with different pools of participants. Moreover in the metamemory literature, metacognitive sensitivity is mainly assessed by Goodman–Kruskal gamma correlations (Kruskal and Goodman 1954) which is a within-subject non-parametric correlation between metacognitive judgements and the accuracy of the memory task. It has been shown that gamma correlations are sensitive to metacognitive bias (i.e. the tendency to be overconfident or underconfident) as well as task performance (Fleming and Lau 2014). Task performance differences are obviously a critical issue in neuropsychological and aging populations, especially where the key theory is that eFOK judgements are less accurate due to deficits in episodic memory (e.g. in Alzheimer’s disease, Ernst et al., 2016). As such, differences in FOKs in the studies of Table 1 are concomitant with a deficit in either episodic recall of recognition (except in autism, see Wojcik et al. 2013). Another solution to evaluate the potential dissociation between eFOK and sFOK is to estimate the cross-task correlations for metacognitive sensitivity in FOK. However, this has rarely been performed as sample sizes are typically small (most of the studies being with patients, see Table 1).

Secondly, investigating the question of whether sFOK and eFOK share common resources relates to the domain generality of metacognition. This assumes that if metacognitive sensitivity depends on domain-general resource, sensitivity indices across two different domains will be correlated. Research to date has exclusively focused on retrospective confidence judgements (RCJs). In contrast to FOKs, RCJs refer to a subjective evaluation of the confidence in a previous decision. As this metacognitive judgement can be performed for decisions in several cognitive domains (e.g. memory, visual perception, and language), it is a good candidate to answer the question of the domain generality of metacognition. As well as an interest in sensitivity, researchers have also examined bias, finding it is domain-general (e.g. Baranski and Petrusic 1995; Ais et al. 2016) that is, people who are over confident on one task will tend to be over confident on another task. The correlation is less clear for sensitivity across tasks. However, under controlled conditions and with appropriate statistical power, it is possible to find cross-domain correlations for metacognitive sensitivity in RCJs (Lee et al. 2018; Mazancieux et al. 2018). In support, a recent meta-analysis of neuroimaging studies has identified a domain-specific neural network involved in RCJs for decision-making and memory tasks (Vaccaro and Fleming 2018).

A further question arises regarding when the metacognitive evaluation is made. As metacognitive judgements can be performed prospectively and retrospectively, it is of interest whether metacognition dissociates according to this variable, especially since the above literature on domain generality tends to focus on retrospective judgements. Several findings support the idea that prospective and retrospective judgements measure distinct aspects of metacognition. They are uncorrelated (e.g. Kelemen et al. 2000), supported by separate brain regions (Chua et al. 2009), and seem to rely on different cues and processes. In the memory field, FOKs are proposed to be based on both the access of partial information of stimuli and the familiarity of the cue (Koriat 1993; Metcalfe et al. 1993; Koriat and Levy-Sadot 2001), whereas RCJs are proposed to be related to the strength of the memory trace (e.g. Yonelinas 1994). Neuroimaging studies support the idea that common and distinct neural mechanisms underpin FOKs and RCJs in memory (Chua et al. 2009). Moreover, studies with neurological populations often exhibit a pattern of impaired sensitivity for prospective judgements and preserved sensitivity for retrospective judgements (e.g. Pannu and Kaszniak 2005). Within the visual perception decision-making field, RCJs are mainly influenced by reaction time and the correctness of the decision, whereas prospective judgements rely more on the judgements made for previous tasks (Fleming et al. 2016). In both literatures, prospective judgement sensitivity is lower than retrospective judgement sensitivity (Carlson 1993; Perfect and Hollins 1996; Fleming et al. 2016). There is some variability in the procedures used and hence in the definition of “prospective.” For instance, in perceptual tasks, the “prospective” judgement is not a prediction before having seen the trial, but are made just before the motor response of the first-order decision, and with all the on-screen information necessary to perform the task. Even so, with such paradigms, there is still significantly lower sensitivity in prospective tasks (e.g. Siedlecka et al. 2016) In contrast to sensitivity, metacognitive bias seems consistent across judgement type (Fleming et al. 2016). FOK judgements, however, have a particular status because they are prospective evaluations but made
Table 1. Review of studies comparing episodic FOK and semantic FOK

| No. | References                  | Participants                                                                 | Metacognitive sensitivity index | Main result—FOK sensitivity                          | Main result—performance                          |
|-----|-----------------------------|------------------------------------------------------------------------------|--------------------------------|------------------------------------------------------|--------------------------------------------------|
| 1   | Bacon et al. (1998)         | 12 placebo 12 low lorazepma dose 12 high lorazepma dose                      | Gamma                          | Lower eFOK for lorazepma groups No difference for sFOK | Lower recall and recognition for both tasks in lorazepma groups |
| 2   | Eakin et al. (2014)         | 50 young adults 56 older adults                                              | Gamma                          | No difference between groups in both FOKs Trend for higher sFOK | Lower episodic recall and recognition for older adults |
| 3   | Morson et al. (2015)        | 35 young adults 16 older adults                                              | Gamma                          | Lower eFOK for older adults No difference for sFOK    | Better semantic recall and recognition for older adults |
| 4   | Pappas et al. (1992)        | 12 older adults 12 patients with AD                                          | Gamma                          | Lower sFOK for AD No difference for eFOK (low for both) | Lower recall for both tasks for AD Lower recall for AD in only episodic memory |
| 5   | Perfect and Hollins (1996)  | 46 young adults                                                               | Gamma                          | Lower for eFOK No difference for sFOK                | No difference in recall Higher recognition for semantic memory Task performance differences not tested |
| 6   | Reggev et al. (2011)        | 23 young adults                                                               | Gamma Hamann                   | No difference between eFOK and sFOK                 | Better semantic and episodic recall and recognition for older adults |
| 7   | Souchay and Smith (2013)    | 16 older adults 16 patients with PD                                           | Gamma                          | Lower sFOK and eFOK for PD                          | Better semantic and episodic recall and recognition for older adults Lower recall and recognition for both tasks for older adults |
| 8   | Souchay et al. (2007)       | 20 young adults 40 older adults                                              | Gamma                          | Lower eFOK for older adults No difference for sFOK    | No difference in task performance between groups |
| 9   | Wojcik et al. (2013)        | 18 children with ASD 18 neurotypical children                                | Gamma                          | Lower eFOK for children with ASD No difference for sFOK |                                                                 |
without a response in the recall phase), participants performed an FOK judgement. More specifically, they had to judge their confidence in recognizing the correct target between two words in a next phase. To do so, they used a 6-point scale ranging from 50% (I will guess the answer) to 100% (very confident in finding the answer). During the last phase, participants performed a 2AFC recognition task where the 40 cues were presented for a second time and participants had to choose between a target and a distractor by pressing either the “s” or the “l” letter. Finally, for each trial, participants had to estimate their level of confidence in their response using the same 6-point scale (ranging from “guessing” to “very confident”).

In the semantic task, participants responded to 40 general-information questions. These questions were word definitions and the participants were asked to recall the word given the definition with a time limit of 15 s. As in the episodic task, in either case, they judged their ability to recognize in a next phase the correct definition between two alternatives. Then in a recognition phase, the definitions were presented a second time and participants had to select the correct answer in a 2AFC task by pressing either the “s” or the “l” letter. They also estimated their level of confidence in their response using the same scale as the episodic memory task.

The stimuli were based on a previous experiment (see Souchay et al. 2007), with the exception that in order to use the meta-\(d’\) framework we presented two alternatives and not four in the recognition phase. We used a feature of the Souchay et al. task, in that the same target words were used in the episodic and the semantic tasks in a counterbalanced fashion (see Fig. 1). Each target has a cue for the episodic memory task and a definition for the semantic memory task. Two lists of 40 targets were created such that each participant was randomly allocated to one set of targets in the episodic condition, and the other in the semantic condition. For both tasks, participants had three training trials in order to familiarize themselves with the task before the test trials. They had 15 s to recall the word before the performing their FOK judgement. For the recognition phase, there was no time limit. The task order was random (on the 92 non-excluded participants, 49 begun with the episodic memory task and 43 begun with the semantic memory task) for each participant and the entire procedure lasted around 45 min.

**Data and statistical analyses**

Our analyses focused on task performance, metacognitive bias, and metacognitive sensitivity. Metacognitive bias was estimated by calculating the difference between mean RCJs or FOKs and mean performance for each participant and each task. Sensitivity was estimated as metacognitive efficiency (meta-\(d’\)/ \(d’\)). Meta-\(d’\) is the \(d’\) that would be expected if the Type 2 receiver operating characteristics (ROC) curve observed for a participant was his or her ideal ROC curve given the task.

---

Figure 1. Summary of the two tasks. The semantic memory task includes two phases and the episodic memory includes three phases.
performance (d’ value). It refers to the signal that is available to perform the Type 2 task (discrimination between correct and incorrect responses). Meta-d’ is in the same unit as d’, thus a ratio can be computed to compare the two. If meta-d’ = d’, participants performed the Type 2 task ideally given their level of Type 1 task performance. A ratio lower than 1 indicates a lower metacognitive sensitivity and a higher ratio a greater metacognitive sensitivity. Several methods exist to estimate meta-d’.

Here, we chose a hierarchical Bayesian model (Fleming 2017) using an extension of the HMeta-d toolbox (https://github.com/metakoglab/HMeta-d) in R software. Two models were used: one for the FOKs and one for the RCJs. Both models estimated a meta-d’/d’ ratio for each participant and each task as well as a group-level parameter for both the episodic and the semantic task and the cross-task correlation. To assess the significance of group-level parameters, we calculated the 95% highest density intervals [HDIs; the smallest interval containing 95% of the Markov chain Monte Carlo (MCMC) samples; Kruschke 2014] on the posterior distributions and looked at its potential overlaps with zero.

The HMeta-d toolbox uses MCMC sampling to estimate posterior distribution over model parameters using JAGS (Plummer 2003) in R (“rjags” package). We modified the HMeta-d code to allow estimation of parameters in R using rjags. As in the HMeta-d toolbox, we discarded early samples of the posterior distributions and ran three chains in order to diagnose convergence problems. We estimated convergence of the three chains with the “potential scale reduction factor” R (Gelman and Rubin 1992). This approach was exactly the same as in Mazancieux et al. (2018) and was carried out according to our preregistration document. Other analyses include analysis of the variance (ANOVA), t-test, and Pearson’s correlations. Outliers were detected using Leverage, RSS, and Cook’s distance. When necessary, Bonferroni corrections were applied.

**Results**

**Type 1 performance**

Task performance was estimated in two ways (Fig. 2). First, we calculated the proportion of correct recall for episodic recall and semantic recall. A recalled item was judged as correct when the exact word was retrieved (we judged as correct those that were recalled in the plural form, even though all targets were singular). Second, we calculated a Type 1 d’ for performance in the recognition task. Paired t-tests showed that performance on the semantic recall (M = 0.27; SD = 0.13) was better than the episodic recall (M = 0.15; SD = 0.09), t(91) = 9.68, P < 0.001, dz = 1.01.

However, the reverse pattern was observed regarding recognition for all items: episodic memory recognition (M = 1.48; SD = 0.68) was better performed than semantic memory recognition (M = 1.23; SD = 0.56), t(91) = 3.10, P = 0.003, dz = 0.32. The pattern of results was the same when comparing proportion of correct recognition only for unrecalled items, t(91) = 5.98, P < 0.001, dz = 0.62, episodic recognition being higher (M = 0.73; SD = 0.11) than semantic recognition (M = 0.65; SD = 0.10).

We were also interested in the intersubject correlations in first-order performance across the two tasks. These analyses revealed positive correlations between episodic and semantic memory for both recall, r = 0.50 [0.32; 0.64], P < 0.001, and recognition for all items, r = 0.26 [0.05; 0.44], P = 0.013, suggesting that participants who performed well on one task also performed well on the other.

**Metacognitive bias**

Metacognitive bias (mean confidence minus mean performance) was calculated for each participant, each task, and for both FOKs and RCJs (Fig. 3). We performed an ANOVA with judgement types and task as factors on metacognitive bias score. It revealed a main effect of task, t(91) = 8.99, P < 0.001,
Metacognitive efficiency

We estimated the group meta-$d’/d’$ ratio for each task and for both FOKs and RCJs (see Fig. 4). To test difference across the four distributions, we performed the difference distribution for each two-by-two comparisons (Table 2). According to overlaps of 95% HDIs with 0, metacognitive efficiency was better for eFOK compared to sFOK. Efficiency for eFOK was however lower than efficiency for both eRCJ and sRCJ. The same pattern was observed for sFOK. Finally, metacognitive efficiencies were the same for eRCJ and sRCJ.

Hierarchical models allow us to estimate correlations between metacognitive efficiency eFOK and sFOK on the one hand and eRCJ and sRCJ on the other hand (Fig. 5). According to overlaps of 95% HDIs with 0, we found no cross-task correlation ($0.22 [-0.89; 0.99]$) for FOKs but a positive correlation ($\rho = 0.47 [0.13; 0.78]$) for RCJs.

Exploratory analyses

Note that the following analyses were not preregistered.

Logistic regressions

Following the idea that FOK can be related to the quality of memory (the memory constraint hypothesis), we were also interested in the relationship between FOK and prior recall. As FOK occurs after a recall attempt, the idea here is to calculate the retrospective sensitivity and bias of the FOK. Therefore, we performed exploratory analyses to see if FOK judgements are related to prior recall. We quantified metacognitive sensitivity using mixed effect logistic regressions, rather than meta-$d’$, as meta-$d’$ assumes that target and distractor distributions are Gaussian with equal variance. Although this is the case for 2AFC recognition, we suggest that this is less applicable for remembered and forgotten words, which limits the use of SDT in this context. Moreover, the estimation of the Type 2 ROC curve from the Type 1 parameters in the meta-$d’$ model as based on the average or response-specific Type 2 ROC curves (i.e. the construction of one curve for “S1” responses for “S2” responses, Maniscalco and Lau 2014).

Consequently, we performed mixed effect logistic regressions to see if FOKs track task performance accuracy for recognition on the one hand and recall on the other hand. This method has been already used as a quantification of metacognitive sensitivity because it is independent from metacognitive bias although not independent from task performance (e.g. Faiivre et al. 2019). We created two models per task: one model in which FOK ratings explain first-order accuracy for the recognition task and a second in which FOK ratings explain first-order accuracy for the recall task. Each model includes confidence as a fixed effect and the estimation of an intercept per participant.

For the episodic memory task, the models reveal that accuracy of recognition was predicted by FOK (estimate $= 0.37, z = 11.70, P < 0.001, \text{OR} = 1.45$) which was also the case for the accuracy of recall (estimate $= 1.65, z = 22.26, P < 0.001, \text{OR} = 5.22$). OR comparison revealed that the effect size was higher for the model in which FOKs explain the accuracy of recall compared to the model in which FOKs explain the accuracy of recognition (OR are 3.6 times larger in the recall model). For the semantic memory task, models reveal that accuracy of recognition was predicted by FOK (estimate $= 2.63, z = 19.89, P < 0.001, \text{OR} = 1.45$) as well as the accuracy of recall (estimate $= 0.97, z = 24.61, P < 0.001, \text{OR} = 2.63$). As for eFOK, the OR comparison shows that the effect size was higher for the model in which FOK explain the accuracy of recall compared to the model in which FOK explain the accuracy of recognition (OR are 2.0 times larger in the recall model).

Correlational analyses

Exploratory analyses revealed that eFOK metacognitive efficiency was correlated with episodic recall, $r = 0.32, P = 0.002$, which was not the case for eRCJ, $r = 0.03, P = 0.781$ (these correlations being significantly different, $z = 2.01, P = 0.036$). Similarly, while sFOK metacognitive efficiency was correlated with semantic recall, $r = 0.33, P = 0.002$, this correlation was significantly not different from the non-significant correlation between semantic recall and sRCJ metacognitive efficiency, $r = 0.09, P = 0.406$ (difference, $z = 1.68, P = 0.093$).
Discussion

The present study focused on the domain-generality of metacognition in two dimensions. First, we wanted to investigate cross-task correlations in episodic and semantic memory indicating a potential general metacognitive resource as in previous works (Lee et al. 2018; Mazancieux et al. 2018; Morales et al. 2018). Second, we wanted to take into account the multifaceted nature of metacognition by comparing two types of metacognitive judgements: FOKs and RCJs. We computed metacognitive efficiency in both the episodic memory task and the semantic memory task separately for FOKs and RCJs. As our tasks contained relatively few trials due to the FOK procedure which requires recall and so cannot use too many items, we estimated metacognitive efficiency in a Bayesian manner (Fleming 2017) and used hierarchical models to estimate cross-task correlations.

First, we reproduced previous findings (Mazancieux et al. 2018) regarding the group-level estimation of the correlation between metacognitive efficiency for episodic and semantic RCJs ($\rho = 0.47 [0.13; 0.78]$ vs. $\rho = 0.41 [0.14; 0.66]$), although the estimation of the current correlation is less precise (with a larger HDI), which can be explained by having roughly half the number of participants in this study ($N = 92$ vs. $N = 181$). This nonetheless suggests that there is a common resource in judging our confidence across episodic and semantic memory. As RCJs are performed after the first-order decision, a suitable candidate for this, especially in memory, is response fluency. That is, we suggest that participants use a common cue from the ease of answering as a basis for gauging the correctness of their response. In support, retrieval or answer fluency (shown as shorter response times) in both general knowledge tasks (Kelley and
Table 2. Means and HDIs of the posteriors of the difference between $\mu$ Mratio distributions for each task pairing

| Difference distributions | FOK episodic memory | FOK semantic memory | RCJ episodic memory | RCJ semantic memory |
|--------------------------|---------------------|---------------------|---------------------|---------------------|
| FOK episodic memory      | 0.20 [0.06, 0.34]   | 0.51 [0.39, 0.64]   | 0.52 [0.39, 0.63]   |
| FOK semantic memory      | 0.71 [0.57, 0.84]   | 0.71 [0.58, 0.84]   |
| RCJ episodic memory      |                     | 0.01 [-0.11, 0.09]  |
| RCJ semantic memory      |                     |                     |

Only the difference distribution between episodic memory and semantic memory overlaps with 0.

Figure 5. Single-subject parameter estimates from the hierarchical model of meta-$d'/d'$ and posterior distributions over $\rho$ in the two covariance matrix determining the correlations between meta-$d'/d'$ across FOKs and RCJs. Distribution of $\rho$ values overlaps with 0 for FOKs (0.22 [−0.89; 0.99]) which is not the case for RCJs (0.47 [0.13; 0.78]).
Lindsay (1993) and episodic recognition tasks (e.g. Benjamin et al. 1998) is associated with higher confidence. Hence, we suggest that retrieval fluency is a diagnostic cue allowing the discrimination between correct and incorrect responses. This cue could even be used beyond the memory domain as answer fluency has been shown to influence confidence in reasoning (Thompson et al. 2013) and response time is negatively correlated with confidence in visual perception decision-making (e.g. Grimaldi et al. 2015).

Second, and most importantly, we found no evidence for a cross-task correlation in eFOK and sFOK. One possible explanation is that even if FOKs imply a prospection of a future recognition task, these judgements are made after a recall test. People can use the output of the retrieval attempt to perform the FOK and therefore use recall as a heuristic for performing FOKs (Schwartz et al. 2016). In contrast with semantic recall, episodic recall involves autonoetic consciousness as a re-experience of the remembered information. Exploratory analyses revealed that for the episodic memory task FOKs metacognitive efficiency was correlated with episodic recall which was not the case for RCJs metacognitive efficiency. This is consistent with the FOK literature showing that FOK judgments are partly based on the recollection process (e.g. Hicks et al. 2002) and especially the retrieval of the encoding context; retrieving information or details about the original encoding context (see the non-criterial recollection hypothesis of episodic FOK; Hertzog et al. 2014; Isingrini et al. 2016). This is also consistent with the classical discrepancy found between impaired episodic FOK and preserved semantic FOK in older adults (e.g. Castel et al. 2016) as recollection is the most impaired process with age (e.g. Clarys et al. 2002), as well as several neurological diseases involving episodic memory impairment (e.g. Alzheimer’s disease, Souchay et al. 2002; multiple sclerosis; Beatty and Monson 1991; patients with frontal lobe lesions, Schnyer et al. 2004). This difference in terms of processes involved in both retrieval mechanisms can also explain why we found a better metacognitive efficiency for eFOK compared to sFOK: partial retrieved information can be used as a cue for performing the FOK judgement eFOK but not sFOK.

A classic distinction in the basis of metacognitive judgements is between experience-based metacognition and information-based metacognition (Koriat and Levy-Sadot 1999). Experience-based metacognition is based on heuristics that are used automatically and that give rise to feeling (as epistemic feelings, see Moulin and Souchay 2014). As an example mentioned above, the fluency heuristic has been shown to influence FOKs (e.g. the cue-familiarity heuristic, Metcalfe et al. 1993). On the contrary, information-based metacognition is based on the application of explicit beliefs or naive theories. Although it is possible that heuristics influence metacognitive judgements across tasks (e.g. the fluency heuristic), we suggest here that beliefs are different in semantic and episodic memory tasks. Semantic memory refers to general knowledge which is shared across people who tend to have an accurate appreciation of what others know (e.g. Juslin 1993). Thus, Koriat (2008) has shown that metacognitive judgements correlate more with the consensual response (the one which is the most chosen by participants) compared to the actual correct response. In episodic memory tasks, such consensus is less likely to occur as retrieval abilities are closely related to the self (e.g. the self-reference effect; Symons and Johnson 1997) and can be thought of as idiosyncratic (Klatzky 1984) and therefore differs across people. Self-referencing has been shown to improve episodic FOK accuracy (Boduroglu et al. 2015).

Going toward the idea that beliefs used for FOKs differ across tasks, Perfect and Hollins (1996) have found a between-subject correlation between FOKs and task performance for semantic memory despite no such relation for episodic memory. We also suggest that explains why sFOK are rarely impaired in patients with memory problems (e.g. Pannu and Kaszniak 2005): information-based cues in sFOK rely more on consensus and do not necessitate autonoesis to be accurate. Semantic FOKs rely more on the inference of what people know or should know based on the activation of a network of related information, whereas episodic FOKs depend on self-knowledge and outputs from the retrieval attempt. That is, if you do not know anything about capitals of African countries, for instance, the semantic FOK will not be able to retrieve any partial information on which to make an accurate FOK judgement: the process is somewhat all-or-nothing. However, in episodic memory, the integration of autonoetic consciousness and the results of the deliberative memory search will likely yield information which is pertinent to episodic FOK judgements. Whilst these inferential processes can occur as well in RCJs, we suggest here that they are less important, as RCJs rely more on the evidence driven by the given response and experience-based metacognition (e.g. answer fluency) therefore explaining the cross-task correlation observed in the present study. Complementing this hypothesis, the exploratory analyses showed that FOK ratings better explains the accuracy of the recall task compared to the accuracy of the recognition. In support, recall was correlated with FOK metacognitive efficiency (albeit in both episodic and semantic tasks).

We therefore here suggest that FOKs are more based on the output of retrieval attempt during the recall (a kind of retrospective recall metacognition) rather than actual prospection of future performance. Thus, the two types of judgement are not correlated because they rely on different types of retrieval process. Future research should experimentally manipulate the semantic and episodic information available at different phases of the task; in order to test the hypothesis that eFOK and sFOK differ according to the cues used, especially because other variables not controlled in this experiment, familiarity, fluency, etc. may bring to bear on the metacognitive decision. Also, here we matched the target word in the two conditions, but it would be of interest to run a task with identical cues (rather than targets) used in the semantic and episodic conditions, such as asking someone to define a word, or retrieve something that was associated with it. It is perhaps possible that our particular pattern of recall and recognition scores have led to the pattern of FOK correlations shown here. Of note we have a higher level of semantic recall, but a lower level of semantic recognition, and there is a large variance in recall scores (see Fig. 2). Whilst the pattern of cross-task correlations in recall but not in FOK, suggests that recall is not a particular concern here, an interesting future experiment would be to manipulate difficulty levels in recall and recognition across the two tasks. Also, FOK sensitivity in the two tasks was correlated to recall.

Finally and regarding metacognitive bias, our results revealed that participants had a lower magnitude of judgement for FOKs compared to RCJs. This is consistent with the fact that less sensory evidence is available for prospective judgements because the task is not yet performed. Therefore, participants are less confident in the FOKs reported here compared to RCJs. We also found that participants tend to give a higher metacognitive judgement for the semantic memory task resulting in more accurate FOKs but an overestimation for RCJs. In accordance with previous work in retrospective judgments (Ais et al. 2016, Mazancieux et al. 2018), we found a cross-task correlation...
for metacognitive bias in FOKs and RCJs. This is congruent with the fact that metacognitive bias is domain-general but also consistent across judgements types (Fleming et al. 2016).

To conclude, this study revealed a cross-task correlation for RCJs in episodic memory and semantic memory suggesting a common resource for metacognitive efficiency in these two tasks. However, no correlation was found across eFOK and sFOK. According to the dual-process view of metacognition (Koriat and Levy-Sadot 1999), we suggest that FOKs and RCJs rely on both experience-based and information-based cues although the amounts of each process differ across judgements. We propose that these processes differ across tasks: sFOK uses inferences based on simply what we know about a subject (noesis) whereas eFOK relies more on inferences based on self-knowledge and an access to the personal past (autonoesis).

Conflict of interest statement. None declared.

References

Ais J, Zylberberg A, Barttfeld P, et al. Individual consistency in the accuracy and distribution of confidence judgments. Cognition 2016;146:377–86.

Bacon E, Danion JM, Kauffmann-Muller F, et al. Consciousness in schizophrenia: a metacognitive approach to semantic memory. Conscious Cogn 2001;10:473–84.

Bacon E, Danion JM, Kauffmann-Muller F, et al. Confidence level and feeling of knowing for episodic and semantic memory: an investigation of lorazepam effects on metamemory. Psychopharmacology 1998;138:318–25.

Baranski JF, Petrusic WM. On the calibration of knowledge and perception. Can J Exp Psychol 1995;49:397–407.

Beatty WW, Monson N. Metamemory in multiple sclerosis. J Clin Exp Neuropsychol 1991;13:309–27.

Benjamin AS, Bjork RA, Schwartz BL. The misinformation of memory: when retrieval fluency is misleading as a metamnemonic index. J Exp Psychol 1998;127:55–68.

Boduroglu A, Pehlivanoglu D, Tekcan AI, et al. Effects of self-referencing on feeling-of-knowing accuracy and recollective experience. Memory 2015;23:736–47.

Carlson BW. The accuracy of future forecasts and past judgments. Organ Behav Hum Decis Process 1993;54:245–76.

Castel AD, Middlebrooks CD, McGillivray S. Monitoring memory in old age: impaired, spared, and aware. In: Dunlosky J and Tauber SK (eds) The Oxford Handbook of Metamemory. New York, NY: Oxford University Press, 2016, 519–35.

Chua EF, Schacter DL, Sperling RA. Neural correlates of metamemory: a comparison of feeling-of-knowing and retrospective confidence judgments. J Cogn Neurosci 2009;21:1751–65.

Clarys D, Isingrini M, Gana K. Mediators of age-related differences in recollective experience in recognition memory. Acta Psychologica 2002;109:315–29.

Eakin DK, Hertzig 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 Cogn 2014;21:27–51.

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

Ernst A, Moulin C, Souchay, C. Anosognosia and metacognition in Alzheimer’s disease: Insights from experimental psychology. In: Dunlosky J and Tauber SK (eds) The Oxford Handbook of Metamemory. Oxford: Oxford University Press, 2016, 451–472.

Faivre N, Vuillaume L, Bernasconi F, et al. Sensorimotor conflicts alter perceptual and action monitoring. Cortex 2019;124:224–34.

Fleming SM. HMeta-d: hierarchical Bayesian estimation of metacognitive efficiency from confidence ratings. Neurosci Conscious 2017;2017: niw018.

Fleming SM, Lau HC. How to measure metacognition. Front Hum 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 Conscious 2016;2016.

Gelman A, Rubin DB. A single series from the gibbs sampler provides a false sense of security. Bayesian Stat 1992;4:625–31.

Grimaldi P, Lau H, Basso MA. There are things that we know that we know, and there are things that we do not know we do not know. Confidence in decision-making. Neurosci Biobehav Rev 2015;55:88–97.

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

Hertzig C, Dunlosky J, Sinclair SM. Episodic feeling-of-knowing resolution derives from the quality of original encoding. Mem Cognit 2010;38:771–784.

Hertzig C, Fulton EK, Sinclair SM, et al. Recalled aspects of original encoding strategies influence episodic feelings of knowing. Mem Cogn 2014;42:126–40.

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

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. Conscious Cogn 2016;41:31–40.

Juslin P. An explanation of the hard-easy effect in studies of realism of confidence in one’s general knowledge. J Exp Psychol Learn Mem Cogn 1990;16:27–51.

Kelley CM, Lindsay DS. Remembering mistaken for knowing: ease of retrieval as a basis for confidence in answers to general knowledge questions. J Mem Lang 1993;32:1–24.

Klatzky RL. Memory and Awareness: An Information-Processing Perspective. New York: WH Freeman, 1984.

Koriat A, Levy-Sadot R. The combined contributions of the cue-ease of retrieval as a basis for confidence in answers to general knowledge questions. J Mem Lang 1993;32:1–24.

Koriat A, Levy-Sadot R. The combined contributions of the cue-familiarity and accessibility heuristics to feelings of knowing. Learn Mem 2001;27:34–53.

Koriat A. How do we know that we know? The accessibility model of the feeling of knowing. Psychol Rev 1993;100:609–39.

Koriat A. Subjective confidence in one’s answers: the consensuality principle. J Exp Psychol 2008;34:945–59.

Koriat A, Levy-Sadot R. Processes underlying metacognitive judgments: information-based and experience-based monitoring of one’s own knowledge. In: Chaien S and Trope Y (eds) Dual-Process Theories in Social Psychology. New York, NY, USA: Guilford Press, 1999, 483–502.

Kruschke J. Doing bayesian data analysis: A tutorial with r, jags, and stan. Academic Press, 2014.

Lee AL, Ruby E, Giles N, et al. Cross-domain association in metacognitive efficiency depends on first-order task types. Front Psychol 2018;9:2464.

Maniscalco B, Lau H. A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. Conscious Cogn 2012;21:422–30.
Maniscalco B, 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, Heidelberg: Springer, 2014, 25–66.

Mazancieux A, Fleming S, Souchay C, et al. Retrospective confidence judgments across tasks: domain-general processes underlying metacognitive accuracy. PsyArXiv. 2018. doi: 10.31234/osf.io/dr7ba.

Mccurdy LY, Maniscalco B, Metcalfe J, et al. Metcalfe J, Schwartz BL, Joaquim SG. The cue-familiarity heuristic in metacognition. Acta Psychol 2013;33:1897–906.

Metcalfe J, Schwartz BL, Joaquim SG. The cue-familiarity heuristic in metacognition. J Exp Psychol 1993;19:851–61.

Morales J, Lau H, Fleming SM. Domain-general and domain-specific patterns of activity supporting metacognition in human prefrontal cortex. J. Neurosci 2018;38:3534–46.

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

Moulin C, Souchay C. Epistemic feelings and memory. In: Perfect TJ and Lindsay D (eds) The Psychology of Learning and Memory: Advances in Research and Theory. London: SAGE Publications Ltd., 2014, 520–38.

Moulin CJ, Souchay C, Smith SJ. Autobiographical memory in Parkinson’s disease: a retrieval deficit. J Neuropsychol 2013;7:164–78.

Souchay C. Metamemory in Alzheimer’s disease. Cortex 2007;43:987–1003.

Souchay C, Isingrini M, Espagnet L. Aging, episodic memory feeling-of-knowing, and frontal functioning. Neuropsychology 2000;14:299–309.

Souchay C, Isingrini M, Gil R. Alzheimer’s disease and feeling-of-knowing in episodic memory. Neuropsychologia 2002;40:2386–96.

Souchay C, Moulin CJ, Clarys D, et al. Diminished episodic memory awareness in older adults: Evidence from feeling-of-knowing and recollection. J Clin Exp Neuropsychol 2006;28:828–40.

Souchay C, Isingrini M, Espagnet L. Aging, episodic memory feeling-of-knowing, and frontal functioning. Neuropsychology 2000;14:299–309.

Souchay C, Isingrini M, Gil R. Alzheimer’s disease and feeling-of-knowing in episodic memory. Neuropsychologia 2002;40:2386–96.

Souchay C, Moulin CJ, Clarys D, et al. Diminished episodic memory awareness in older adults: Evidence from feeling-of-knowing and recollection. J Clin Exp Neuropsychol 2006;28:828–40.

Symons CS, Johnson BT. The self-reference effect in memory: a meta-analysis. Psychol Bull 1997;121:371–94.

Thompson VA, Turner JAP, Pennycook G, et al. The role of answer fluency and perceptual fluency as metacognitive cues for initiating analytic thinking. Cognition 2013;128:237–51.

Vaccaro AG, Fleming SM. Thinking about thinking: a coordinate-based meta-analysis of neuroimaging studies of metacognitive judgments. Brain Neurosci Adv 2018;2:2398212818810591.

Wojcik DZ, Moulin CJ, Souchay C. Metamemory in children with autism: exploring “feeling-of-knowing” in episodic and semantic memory. Neuropsychology 2013;27:19–27.

Yonelinas AP. Receiver-operating characteristics in recognition memory: evidence for a dual-process model. J Exp Psychol 1994;20:1341–54.