Metacognition Learning (2016) 11:305–316
DOI 10.1007/s11409-015-9152-4

Methodological concerns: the feeling-of-knowing task affects resolution

Bennett L. Schwartz[1] · **Aysecan Boduroglu**[2] · **Ali İ. Tekcan**[2]

Received: 16 July 2015 / Accepted: 25 November 2015 / Published online: 4 December 2015

Abstract In traditional feeling-of-knowing procedures, participants make judgments on unrecalled items only (e.g. Hart 1965). However, many researchers elicit feeling-of-knowing judgments (FOKs) on all items. When FOKs are made on all items, participants may use recall as a basis for judgments, leading to higher magnitude judgments for recalled items, but causing a relative floor effect for judgments for unrecalled items. We suspected that resolution (relative accuracy) would be better when FOKs are made on all items than when they are made on unrecalled items only. We examined the issue by comparing across studies, reanalyzing data from another experiment, and by conducting an original experiment. In the literature review, we included 83 conditions across 52 studies. We found that feeling-of-knowing judgments made on all items showed higher resolution than feeling-of-knowing judgments made on unrecalled items. This was replicated in the reanalysis of existing data of a single study that used both methods. In the original experiment, we collected feeling-of-knowing judgments for general-information questions. The experiment confirmed that resolution for predicting recognition was higher when feeling-of-knowing judgments were made on all items than when they were made only on unrecalled items. We discuss both methodological and theoretical implications of these data.

Keywords Feeling-of-knowing · Metamemory · Methodology

Metamemory concerns the monitoring and control of memory processes (Nelson and Narens 1980; Metcalfe 2000). A common way in which metamemory has been studied is by examining feeling-of-knowing judgments (henceforth, FOKs). FOKs typically correspond to judgments of the future ability to recognize currently-unrecalled items (Hart 1965; Souchay 2013). The typical procedure, known as RJR for recall-judgment-recognition, involves a recall

✉ Bennett L. Schwartz
schwartb@fiu.edu

[1] Department of Psychology, Florida International University, University Park, Miami, FL 33199, USA

[2] Boğaziçi University, Istanbul, Turkey

stage, followed by a stage in which participants make FOKs on unrecalled items, followed by a recognition test. Decades of research have demonstrated that FOKs accurately predict the likelihood of subsequent recognition of unrecalled targets (Koriat 2000; Souchay 2013). Thus, FOKs illustrate a paradoxical state in which a person reports confidence regarding retrieval of a target item that cannot be immediately accessed, and this has captured researchers' attention because it has the potential to reveal insights about the nature of conscious recollection (Koriat 2000).

Some researchers have found the above methodology insufficient, particularly when investigating the mechanisms that underlie FOK. For example, examining only unrecalled items can create a different distribution and total number of items in different experimental conditions, thus rendering comparisons across conditions difficult. To correct these problems, researchers modified the traditional recall-judgment-recognition procedure to include FOKs on all items, not just those that are unrecalled (e.g., Boduroglu et al. 2015, 2014; Eakin and Hertzog 2012; Koriat 1993, 1995; Souchay & Isingrini 2012). We suspect that, in this methodology, participants employ whether they have just recalled the target a basis for their FOK. It is likely that participants implicitly assume that if they just recalled the target, there is a strong chance that they will recognize it, and the FOK will assume the maximum value when an item has been recalled (see Nelson et al. 2004). However, it is possible that participants may anticipate forgetting or anticipate being confused by recognition choices, thus leading to lower FOKs for recalled items (Kornell 2012). Thus, if participants employ recall as a basis for FOK is an empirical question.

Requiring that participants make FOKs on all items has the following advantages over the traditional method of collecting FOKs on unrecalled items. First, requiring FOK on all items eliminates selection effects on FOK that occur in the traditional method because there are usually systematic differences between items recalled and those not recalled (Bauml 2008; Mickes et al. 2013; Schwartz and Metcalfe 1994). If the researcher is comparing the effects of two or more levels of a variable on FOK, collecting FOK on all items will allow causal inferences to be made if there is a difference among conditions. Second, participants may not know if their answer is correct or incorrect (Koriat 1993). If participants output an answer but think it is incorrect, this may be reflected in FOKs. If FOK is collected only on unrecalled items, this cannot be addressed (Boduroglu et al. 2015). For these reasons, much of the recent literature on FOKs collects judgments on all items (see Appendix).

The issue of whether FOKs are made on all items or only unrecalled items is relevant to the discussion of the heuristics underlying FOKs. Most researchers agree that FOKs are partially based on unconscious inferences that participants make to judge the likelihood of future recognition (Koriat 1993; Metcalfe 1993). though conscious reasoning may also play a role. Research addressing the mechanisms underlying FOKs has focused on the retrieved partial target information, related information, and cue familiarity (Brewer et al. 2010; Hertzog et al. 2014; Thomas et al. 2012). We suspect that when people make FOKs on all items, an important heuristic will be whether they can recall the target itself. This heuristic will likely overwhelm the heuristics used when evaluating only unrecalled items, such as partial retrieval and cue familiarity. Indeed, Boduroglu et al. (2015) reported much higher FOK magnitude in an experiment using the all-items procedure than in an experiment using the traditional procedure. This change in the heuristics used to generate FOKs may cause differences in resolution (relative accuracy). Because successful recall is likely to be highly predictive of successful recognition, resolution should be particularly high in the all-items condition. That is, successful recall may be a better predictor of later recognition than is the retrieval of partial information or

cue familiarity, which influence FOKs in the traditional method (see Schwartz et al. 2014). In the current study, we do not experimentally address the issue of mechanism or how the mechanisms changes from FOKs on all items to FOKs on unrecalled items, but we ask how the differences in methodology affect resolution. It is our view that having FOKs made on all items may shift participants to base their FOKs on whether they have recalled the item, whereas making FOKs on only unrecalled items will more likely to be based on other heuristics. If true, one might expect differences in the resolution or relative accuracy of each kind of FOK.

Typically in any memory test, some of the target items will be recalled, whereas others will not be. When FOKs are made on all items, we expect FOK for unrecalled items to be made on a restricted range because the high end of the FOK scale will mostly likely be devoted to already retrieved targets. If any unrecalled items are recognized, as will likely be the case, this will lower the resolution score. It also may be that the restricted range may compress the scale used by participants without affecting relative order, in which case one might not see differences in FOK between unrecalled items in the two paradigms. Our actual results show evidence that for unrecalled items when binary FOK scales are used, the restricted range does lower resolution, but not when Likert or percent confidence scales are used.

Although our paper focuses on FOKs, there is a related literature on the effects of recall on judgments of learning. In research on judgments of learning, metacognitive resolution improves when participants have an opportunity to attempt retrieval before making judgments of learning. In this procedure, known as the PRAM methodology, there is an initial recall stage, followed by a judgment stage, rather than combining the two (Nelson et al. 2004; Narens et al. 2008). In this methodology, the resolution of judgments of learning is higher because the pre-judgment recall increases the salience of retrieved information. Indeed, in the PRAM methodology, the JOLs made are very similar to the FOKs in the all-item condition. Similarly, we suspect that when FOKs are made after recall, the salience of the recalled information will lead to a boost in resolution.

Our goal in this paper is to examine how FOK methodology (FOKs on all items vs. only the unrecalled items) affects the resolution of FOKs. We offer two hypotheses. Our first hypothesis is based on the idea that when FOKs are made on all items, recall will be a reliable heuristic for predicting future recognition (e.g., Koriat and Levy-Sadot 2001). That is, just-recalled likely means to-be-recognized. Because subsequent correct recognition is likely to be highly correlated with recall, when participants rely on this heuristic, their resolution will be higher than when they must rely only on less effective heuristics, such as partial information or cue familiarity, as is the case when FOKs are only made on unrecalled items. Thus, if all else is equal, FOK resolution will be higher when FOKs are obtained on all items than when they are only obtained on unrecalled items. Although this hypothesis may appear obvious, no research to date has addressed it.

Our second hypothesis concerns only the resolution of unrecalled items. Because making FOKs on all items will cause high judgments to be reserved for recalled items (Boduroglu et al. 2015). there will be a restricted range for unrecalled items when all items are given FOKs. That is, in the all-items condition, unrecalled items will tend to receive lower FOKs than those same items would in the unrecalled-only condition. In other words, people will implicitly save high judgments for recalled items in the all-items condition, but then use those same higher judgments when only making FOKs on unrecalled items. Thus, FOK resolution for unrecalled items will be lower when we compare unrecalled items from an experiment in which FOKs were collected on all items to unrecalled items from an experiment in which FOKs were collected on only unrecalled items.

We tested the hypotheses in three ways. First, we conducted a survey of the existing FOK literature. We divided the literature into those that compared FOK resolution on all items (alternate approach; all-items condition), on unrecalled items within the all-items condition (alternate approach; unrecalled-all condition) and studies in which FOK was measured only for unrecalled items (traditional approach; unrecalled-only condition). We then measured the mean gamma correlation across these three conditions across all of the studies we compiled. Second, we did an original experiment, in which we compared the resolution of FOKs when participants made FOKs in an all-items condition and when participants made FOKs in an unrecalled-only condition using a binary FOK scale. Third, we re-examined data from an earlier experiment in which conditions included both FOKs on all items and FOKs on unrecalled items using a continuous FOK scale (Boduroglu et al. 2015). We predicted that FOK resolution, as measured by gamma correlations will be greater in the all-items condition than the unrecalled-only condition. We also predicted that the gamma correlation will be higher for unrecalled items in the unrecalled-only condition than the unrecalled-all condition (see Barrett et al. 2014, Masson and Rotello 2009, and Nelson 1984 for discussions of the use of gamma as a measure of resolution).

Cross-study analysis

The studies that met our criteria and were included in the analysis are provided in Appendix. The following criteria were used to determine if a study qualified for the analysis. First, we required the instructions to participants concern judgments or predictions of future recognition performance, that is, FOK instructions. This excluded studies that asked for predictions of recall or predictions of knowing without specifying the test. Second, we focused on younger adults and excluded conditions that tested older adults because there may be more variance among older populations; younger adults across studies tend to be college undergraduates, and we considered them to be a uniform sample. Third, we excluded all studies looking at children participants (< age 18) for similar reasons. Fourth, we excluded special populations, such as psychiatric patients or neuropsychological patients as some of these groups may show metacognitive deficits (e.g. Baran et al. 2009; Souchay 2013). We used the following inclusion criteria. In order to have one common metric to compare across studies, we only included studies in which FOK resolution was measured using gamma correlations. Some newer studies use signal-detection analysis, as per recommendations of Masson and Rotello (2009). and some older studies simply report the percent correct for FOK items at each level of FOK. In the all-items condition, we included studies in which FOK was solicited before recall and when FOK was solicited after recall (see Appendix for the studies used). Although there may be differences between these procedures, we grouped them together because we anticipate that generating an answer will serve as a major basis for FOK in either case. We included studies that used either semantic-memory materials (such as general-information questions) or episodic-memory materials (such as recently-learned paired associates), although it may be that some differences exist between these materials with respect to accuracy (Bacon and Izaute 2009).

The final sample consisted of 52 papers reporting 83 separate conditions, each reporting a gamma correlation between FOK and recognition. Of these conditions, 39 were unrecalled-only condition, 25 were in the all-items condition, and 19 conditions measured FOK for unrecalled items in studies that collected FOK on all items.

The results showed effects in the expected directions, F (2, 80) = 17.12, $p < 0.01$, MsE = 0.34, η^2 = 0.30. (see Fig. 1). Tukey HSD post-hoc tests confirmed that the mean gamma correlation for FOKs in the all-items condition (M = 0.57, SD = 0.10) was higher than the unrecalled-only condition (M = 0.40, SD =. 16), and the unrecalled-all condition (M = 0.34, SD = 0.16). There was not a significant difference between the FOKs in the unrecalled-only condition and the unrecalled-all condition. Thus, the first hypothesis that resolution will be higher in the all-items condition than in the unrecalled-only condition was supported (but our second hypothesis was not). This supports the contention that participants are using recall as an effective predictor of later recognition.

Experiment

Participants

The participants were 55 Florida International University students who received course credit for participating. There were 27 participants in the all-items condition, and 28 participants in the unrecalled-only condition. Each was tested individually on a Macintosh computer during a session that lasted approximately ½ hour. The participants in the all-items condition were run during the winter semester of 2012, whereas the participants in the only-unrecalled condition were run during the fall semester of 2011. The Florida International University Internal Review Board approved these studies.

Materials

Superlab4 software was used to run the experiments and collect the data. The stimuli were 79 general information questions (e.g., "For which country is the rupee the monetary unit? India), taken from the Nelson–Narens norms (Nelson and Narens 1980).

Procedure

Participants were first given detailed instructions about the procedure. They were told that they would be answering a series of general-information questions. Participants were encouraged not to guess and to indicate that they did not know, rather than to guess on any particular trial.

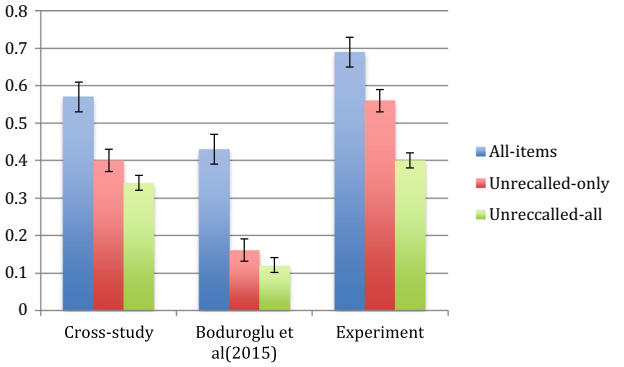

Fig. 1 Gamma correlations between FOK and recognition as a function of whether FOKs were collected on all items or only the unrecalled items. Gammas are highest in the all-items condition. Y-axis represents gamma correlations. Standard error bars included

Participants were instructed that they could pace themselves as they answered questions. They were given the following instructions concerning feeling-of-knowing judgments. A FOK state was defined as follows: "A feeling of knowing means that you feel that you think that you will correctly recognize the answer when you see it among a list of eight alternatives." They were not given instructions on the differences between FOKs on all items and on unrecalled items, as this was a between-subjects variable.

The order of presentation of general information questions was randomized for each participant. Each question appeared in the middle of the screen, and a prompt appeared beneath the question. The question remained on the screen until the participants spoke their response, and the experimenter typed in the answer (to avoid spelling mistakes). In the all-items condition, participants made a FOK on the item, regardless of whether the response had been a "do not know" response or an answer – correct or incorrect. In the unrecalled-only condition, if the participant typed in the correct response, they moved on to the next question. Commission errors were not followed by FOKs in the unrecalled-only condition. As a consequence, we also eliminated commission errors from consideration in the all-items condition as well. In both conditions, the prompt was ""FOK?" The participants typed in a "Y" if they were experiencing a positive FOK and an "N" if they were not. The FOKs were dichotomous – yes or no – in this study because it was originally run as a control condition for a study measuring tip-of-the-tongue states and feelings-of-forgetting, the former of which is normally a dichotomous judgment (Schwartz 2006).

After the participants attempted recall for all 79 questions, they were given a recognition test for all questions. They were again shown the question, followed by eight alternatives, one of which was the correct answer (Wilkinson and Nelson 1984). A number accompanied each alternative. The participants typed in the number associated with the answer that they thought was correct. The recognition distractors were all close associates of the correct answer. For example, for the question "What was the last name of the first person on the moon?" all of the distractors were other prominent astronauts from the same time period. At the end of the session, the participants were debriefed and were given credit in their introductory psychology course.

Results

Recall in the all-items condition (36 %) and in the unrecalled condition (37 %) was not different, $F < 1$. Correct recognition of unrecalled answers (55 %) also did not vary across conditions, $F < 1$. FOKs were made on 66 % of all trials in the all-items condition and 63 % of all unrecalled trials in the unrecalled-only condition, $F < 1$. We compared the number of positive FOKs for unrecalled items (that is, though items for which people indicated that they had a FOK) across the all-items and unrecalled conditions. There were more positive FOKs for unrecalled trials in the unrecalled-only condition (63 %) than in the unrecalled-all condition (53 %), $F\ (1, 54) = 7.44$, $MsE = 0.35$, $p = 0.01$, $\eta^2 = 0.14$. In the all-items condition, the gamma correlation between recall and FOKs (0.76) was significantly above zero, $F\ (1, 26) = 29.13$, $p < 0.01$, $\eta^2 = 0.29$. This correlation suggests that recall was used as an index of likely recognition, in keeping with the hypotheses.

In order to determine gamma correlations between FOK and recognition, we needed to ensure that participants used both levels of FOK and had recognition answers both correct and incorrect. Two participants (one in each condition) had indeterminate gammas because in one case, the participant never made a positive FOK, and in the other condition, the participant did

not have any correct on the recognition test. Therefore the analyses are based on 26 participants in the all-items condition and 27 in the unrecalled-only condition. The mean gamma correlation was higher (M = 0.69, SD = 0.24) in the all-items condition than it was in the unrecalled-only condition (M = 0.56, SD = 0.16), F (1, 51) = 5.46, MsE = 0.23, p = 0.02, η^2 = 0.10. Thus, the experiment supports our hypothesis that gamma correlations will be higher in the all-items condition.

We also compared the gamma correlation for unrecalled items in the all-items condition with the gamma correlations for the unrecalled-only condition. The mean gamma for the unrecalled-all items was (0.41, SD = 0.30), significantly lower than the gamma correlation in the unrecalled-only condition (0.56, SD = 0.16), F (1,51) = 4.4, MsE = 0.26, p = 0.04, η^2 = 0.08.

Re-analysis of Boduroglu et al. (2015)

Boduroglu et al. (2015) were interested in the effects of self-referencing on the magnitude and resolution of FOKs. They examined this by comparing memory and FOK performance across self-reference and semantic-control conditions, using recently learned word pairs. FOKs were made on a 0 to 100 scale, with 0 indicating that participants definitely would not recognize the target whereas 100 represented strong confidence that the target would be recognized (see Boduroglu et al. 2015 for more details on the methodology). Most relevant here is that Boduroglu et al.'s main experiment used the all-items condition, but they included a secondary experiment that looked at the effects of self-referencing on FOKs in an unrecalled-only condition using exactly the same materials. In the following analysis, we examined the effect on resolution across the all-items condition and the unrecalled-only condition.

The results showed a strong effect in the expected directions, F (2, 154) = 17.57, MsE = 0.17.7, $p < 0.01$, η^2 = 0.19, for the self-reference condition. Tukey HSD post-hoc tests confirmed that the mean gamma correlation for FOKs in the all-items condition (M = 0.48, SD = 0.22) was higher than the unrecalled-only condition (M = 0.25, SD =. 40), and the unrecalled-all condition (M = 0.12, SD = 0.39). There was not a significant difference between the FOKs in the unrecalled-only condition and the unrecalled-all condition. The pattern was identical for the semantic-encoding condition, F (2, 154) = 11.08, $p < 0.01$, MsE = 19.9, η^2 = 0.13. Tukey HSD post-hoc tests confirmed that the mean gamma correlation for FOKs in the all-items condition (M = 0.37, SD = 0.31) was higher than the unrecalled-only condition (M = 0.06, SD = 0.37), and the unrecalled-all condition (M = 0.12, SD = 0.40). Figure 1 shows the data averaged across both self and semantic conditions. There was not a significant difference between the FOKs in the unrecalled-only condition and the unrecalled-all condition. Thus, like the cross-study analysis, these data only support the hypothesis that resolution will be higher in the all-items condition than the unrecalled-only condition. The data from Boduroglu et al. also support the idea that recall is source of information for determining the FOKs. FOKs for recalled items were 92.00 % across conditions, whereas FOKs for unrecalled items in the all-recall condition were 44.74 % (see Fig. 1 in Boduroglu et al. 2015).

General discussion

We found that FOK resolution varies depending on the method used to collect FOKs. FOK resolution is higher when FOKs are made on all items than when made only unrecalled items.

We found this pattern in an analysis across existing studies, in the re-analysis of the Boduroglu et al. (2015) data, and in our original experiment, thereby confirming our first hypothesis. Higher resolution here only implies that the all-items methodology results in higher gamma correlations. We conjecture that it is likely that this higher resolution results because participants can employ a recall-based strategy, that is, if recalled-then-likely to be recognized. If this is so, then the improvement in resolution is a function of enhanced metacognition in the all-items condition and not simply a methodological artifact. However, we found little support for our second hypothesis that gamma correlations would be higher in the unrecalled-only condition than in the unrecalled-all condition.

This current finding that gamma correlations are higher in the all-items condition than the unrecalled-only condition presents several challenges for researchers. First, the findings here suggest that researchers should use caution when comparing FOK accuracy across studies. It is possible that one lab finds higher accuracy not because of specific differences in their variables or population but because they used a different methodology to obtain those FOKs. Thus, researchers need to account for methodology when making comparisons about FOK resolution.

Each methodology has advantages for asking different questions. For example, if the interest is in examining differences across conditions, such as when examining mechanisms of FOK, then using the all-items approach is preferable, as one can keep the different conditions equivalent except for the experimental variables (Koriat 1995). In contrast, if the interest is in examining the nature of interrupted or blocked retrieval, the unrecalled-only approach is preferable, as this methodology allows one to focus on retrieval failure (Brown 2012).

The current results suggest differences in the processes underlying all-items FOKs and unrecalled-only FOKs. We suspect that successful recall is used as a cue for determining FOK when the all-items condition is used. In the all-items condition, a successful heuristic is "if I recalled it, then I will recognize it later." Indeed, there was a strong correlation between successful recall and the likelihood of indicating an FOK in the all-items condition, consistent with this hypothesis. In the re-analysis of Boduroglu et al. (2015). we also found much higher FOKs when items had been recalled then when they had not in the all items condition. When this heuristic is available to participants, they may weight it more heavily than such heuristics as cue familiarity and partial target information (e.g., Brewer et al. 2010; Thomas et al. 2012). Indeed, Thomas et al. (2011, 2012). found that recalling attributes of the target stimulus – without recalling the stimulus name – led to higher FOKs. Boduroglu et al. (2015) also found higher FOK magnitude in the all-item condition than the unrecalled-only condition.

The recall heuristic is consistent with the research on FOKs that suggest that the basis FOKs are unconscious inferences made at the time of judgment (Brewer et al. 2010; Hertzog et al. 2010; Koriat 1993, 1995; Metcalfe et al. 1993). With respect to FOKs on unrecalled items, the research has focused on partial retrieval of target information, the attempted retrieval of related information, and the familiarity of the cue. However, when people are making FOKs on items they just recalled, they have an even bigger potential source of information for their judgment – that is, whether or not they just recalled the target.

An unanswered question is how the difference between all-items FOKs and unrecalled-only FOKs affects the magnitude of FOKs. Our data are consistent with the view that when FOKs are made on all items, the FOKs for already-recalled items will be high, but this procedure may

also lower FOKs for unrecalled items. Indeed, this is the pattern recently observed across-experiments by Boduroglu et al. (2015). Further study will be required to determine how the different methodologies affect FOK magnitude.

Our conclusion is that the way FOK judgments are elicited may lead to important differences in FOK resolution measure. Thus, we urge caution when comparing FOKs made with the all-items procedure to those made with the unrecalled-only procedure. We assert that both yield valuable information and inform us about the nature of metacognition, but that the change of procedure induces a number of differences that must be accounted for. In this paper, we enumerate one of them – namely that FOK resolution is affected by whether FOKs are made on all items or just unrecalled items. FOK resolution is higher in the all-items condition than the unrecalled-only condition, suggesting that participants use recall as a heuristic for determining FOK.

Appendix

Studies using FOK on unrecalled only

Boduroglu et al. (2015)
Hicks and Marsh (2002)
Hosey et al. (2009)
Korenman and Peynircioglu (2004)
Le Berre et al. (2010)
Metcalfe et al. (1993)
Maril et al. (2003)
Peynircioglu et al. (1998)
Peynircioglu and Tekcan (2000)
Pinon et al. (2005)
Rabinovitz and Peynircioğlu (2011)
Schwartz et al. (2014)
Tekcan et al. (2007)
Tekcan and Akturk (2001)
Thomas et al. (2011)
Tuna et al. (2005)
Watier and Collin (2011)
Widner et al. (1996)

Studies using all-items but analyzing unrecalled only

Eakin and Hertzog (2012)
Eakin et al. (2014)
Hertzog et al. (2010)
Hertzog et al. (2014)
MacLaverty and Hertzog (2009)
Sacher et al. (2009)
Sacher et al. (2013)
Souchay et al. (2007)

Studies using all-items

Boduroglu et al. (2014)
Boduroglu et al. (2015)
Eakin and Hertzog (2012)
Koriat (1993)
Koriat (1995)
Koriat and Levy-Sadot (2001)
MacLaverty and Hertzog (2009)
Modirrousta and Fellows (2008)
Perrotin et al. (2006)
Perrotin et al. (2008). Reggev et al. (2011)
Sacher et al. (2014)
Souchay & Isingrini (2012)
Swerts and Krahmer (2005)
Watier and Collin (2011)
Watier and Collin (2012)

References

Bacon, E., & Izaute, M. (2009). Metacognition in schizophrenia: processes underlying patients' reflections on their own episodic memory. *Biological Psychiatry*, *66*, 1031–1037.

Baran, B., Tekcan, A. İ., Gürvit, H., & Boduroglu, A. (2009). Episodic memory and metamemory in Parkinson's disease patients. *Neuropsychology*, *23*(6), 736.

Barrett, A., Dienes, Z., & Seth, A. K. (2014). Measures of metacognition on signal- detection theoretic models. *Psychological Methods*, *18*, 535–552.

Bauml, K. H. (2008) Inhibitory processes. In H. L. Roediger, III (Ed.), Cognitive Psychology of Memory. Vol. [2] of Learning and Memory: A Comprehensive Reference, 4 vols. (J. Byrne Editor), pp. [195–220] Oxford: Elsevier.

Boduroglu, A., Tekcan, A.İ., & Kapucu, A. (2014). The relationship between executive functions, episodic feeling-of-knowing, and confidence judgments. *Journal of Cognitive Psychology*, *26*, 333–345. http://dx.doi.org/10.1080/20445911.2014.891596.

Boduroglu, A., Pehlivanoglu, D., Tekcan, A. İ., & Kapucu, A. (2015). Effects of self- referencing on FOK accuracy and recollective experience. *Memory*, *23*, 736–747. http://dx.doi.org/10.1080/09658211.2014.925927

Brewer, G. A., Marsh, R. L., Clark-Foos, A., & Meeks, J. T. (2010). Noncriterial recollection influences metacognitive monitoring and control processes. *The Quarterly Journal of Experimental Psychology*, *63*, 1936–1942.

Brown, A. S. (2012). *Tip of the tongue states*. New York: Psychology Press.

Eakin, D. K., & Hertzog, C. (2012). Age invariance in feeling of knowing during implicit interference effects. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *67*(5), 555–562.

Eakin, D. K., Hertzog, C., & Harris, W. (2014). Age invariance in semantic and episodic metamemory: both younger and older adults provide accurate feeling-of-knowing for names of faces. *Aging, Neuropsychology, and Cognition, 21*, 27–51.

Hart, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational Psychology*, *56*, 208–216.

Hertzog, C., Dunlosky, J., & Sinclair, S. M. (2010). Episodic feeling-of-knowing resolution derives from the quality of original encoding. *Memory & Cognition*, *38*, 771–784.

Hertzog, C., Fulton, E. K., Sinclair, S. M., & Dunlosky, J. (2014). Recalled aspects of original encoding strategies influence episodic feeling of knowing. *Memory & Cognition*, *42*, 126–140.

Hicks, J. L., & Marsh, R. L. (2002). On predicting the future states of awareness for recognition of unrecallable items. *Memory & Cognition*, *30*, 60–66.

Hosey, L. A., Peynircioğlu, Z. F., & Rabinovitz, B. E. (2009). Feeling of knowing for names in response to faces. *Acta Psychologica*, *130*, 214–224.
Korenman, L. A., & Peynircioglu, Z. F. (2004). The role of familiarity in episodic memory and metamemory for music. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 917–922.
Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review*, *100*, 609–639.
Koriat, A. (1995). Dissociating knowing and the feeling of knowing: further evidence for the accessibility model. *Journal of Experimental Psychology: General*, *124*(3), 311–333.
Koriat, A. (2000). The feeling of knowing: some metatheoretical implications for consciousness and control. *Consciousness and Cognition*, *9*, 149–171.
Koriat, A., & Levy-Sadot, R. (2001). The combined contributions of the cue-familiarity and accessibility heuristics to feelings of knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(1), 34–53.
Kornell, N. (2012). A stability bias in human memory. In N. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 4–7). New York: Springer.
Le Berre, A.-P., Pinon, K., Vabret, F., Pitel, A.-L., Allain, P., Eustache, F., et al. (2010). Study of metamemory in patients with chronic alcoholism using a feeling-of-knowing episodic memory task. *Alcoholism: Clinical and Experimental Research*, *34*, 1888–1898.
MacLaverty, S. N., & Hertzog, C. (2009). Do age-related differences in episodic feeling of knowing accuracy depend on the timing of the judgement? *Memory*, *17*, 860–873.
Maril, A., Simons, J. S., Mitchell, J. P., Schwartz, B. L., & Schacter, D. L. (2003). Feeling of knowing in episodic memory: an event-related fMRI study. *NeuroImage*, *18*, 827–836.
Masson, M. E. J., & Rotello, C. M. (2009). Sources of bias in the Goodman-Kruskal gamma coefficient measure of association: implications for studies of metacognitive processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 509–527.
Metcalfe, J. (1993). Novelty monitoring, metacognition and control in a compositeholographic associative recall model: interpretations for Korsakoff amnesia. *Psychological Review*, *100*, 3–22.
Metcalfe, J. (2000). Metamemory: theory and data. In E. Tulving & F.I.M. Craik (Eds.), *Oxford handbook of memory* (pp. 197–211). London: Oxford University Press.
Metcalfe, J., Schwartz, B. L., & Joaquim, S. G. (1993). The cue familiarity heuristic in metacognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 851–861.
Mickes, L., Seale-Carlisle, T. M., & Wixted, J. T. (2013). Rethinking familiarity: remember/know judgments in free recall. *Journal of Memory and Language*, *68*, 333–349.
Modirrousta, M., & Fellows, L. K. (2008). Medial prefrontal cortex plays a critical and selective role in "feeling of knowing" meta-memory judgments. *Neuropsychologia*, *46*, 2958–2965.
Narens, L., Nelson, T. O., & Scheck, P. (2008). Memory monitoring and the delayed JOL effect. In J. Dunlosky, & R. A. Bjork (Eds.), *Handbook of metamemory and memory* (pp. 137–154). New York: Psychology Press.
Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of- knowing predictions. *Psychological Bulletin*, *95*, 109–133.
Nelson, T. O., & Narens, L. (1980). Norms of 300 general-information questions: accuracy of recall, latency of recall, and feeling-of- knowing ratings. *Journal of Verbal Learning & Verbal Behavior*, *19*, 338–368.
Nelson, T. O., Narens, L., & Dunlosky, J. (2004). A revised methodology for research on metamemory: Pre-judgment recall and monitoring (PRAM). *Psychological Methods*, *9*, 53–69.
Perrotin, A., Isingrini, M., Souchay, C., Clarys, D., & Taconnat, L. (2006). Episodic feeling of knowing accuracy and cued recall in the elderly: evidence for double dissociation involving executive functioning and processing speed. *Acta Psychologica*, *122*, 58–73.
Perrotin, A., Tournelle, L., & Isingrini, M. (2008). Executive functioning and memory as potential mediators of the episodic feeling-of-knowing accuracy. *Brain and Cognition*, *67*, 76–87.
Peynircioglu, Z. F., & Tekcan, A. İ. (2000). Feeling of knowing for translations of words. *Journal of Memory and Language*, *43*, 135–148.
Peynircioglu, Z. F., Tekcan, A. İ., Wagner, J. L., Baxter, T. B., & Shaffer, S. D. (1998). Name or hum that tune: feeling of knowing for music. *Memory and Cognition*, *26*, 1131–1137.
Pinon, K., Allain, P., Kefi, M. Z., Dubas, F., & Le Gall, D. (2005). Monitoring processes and metamemory experience in patients with dysexecutive syndrome. *Brain and Cognition*, *57*, 185–188.
Rabinovitz, B. E., & Peynircioğlu, Z. F. (2011). Feeling-of-knowing for songs and instrumental music. *Acta Psychologica*, *138*, 74–84.
Reggev, N., Zuckerman, M., & Maril, A. (2011). Are all judgments created equal? An fMRI study of semantic and episodic metamemory predic- tions. *Neuropsychologia*, *49*, 1332–1342.

Sacher, M., Taconnat, L., & Souchay, C. (2009). Divided attention at encoding: effect on feeling-of-knowing. *Consciousness & Cognition*, *18*, 754–761.
Sacher, M., Isingrini, M., & Taconnat, L. (2013). Effects of aging and divided attention on episodic feeling-of-knowing accuracy. *Acta Psychologica, 144*, 258–263.
Sacher, M., Landré, L., & Taconnat, L. (2014). Age-related differences in episodic feeling-of-knowing arise from differences in memory performance. *Memory*. doi:10.1080/09658211.2013.870210.
Schwartz, B. L. (2006). Tip-of-the-tongue states as metacognition. *Metacognition and Learning*, *1*, 149–158.
Schwartz, B. L., & Metcalfe, J. (1994). Methodological problems and pitfalls in the study of human metacognition. In J. Metcalfe & A. Shimamura (Eds.) Metacognition: Knowing about knowing. pp. 93–114.
Schwartz, B. L., Bacon, E., & Pillot, M. (2014). The contribution of partial and contextual information to the feeling of knowing in episodic memory. *Consciousness & Cognition*, *29*, 96–104.
Souchay, C. (2013). Métamémoire et troubles de la mémoire : l'exemple du feeling-of knowing. *Review de Neuropsychologie*, *6*, 265–272.
Souchay, C., & Isingrini, M. (2012). Are feeling-of-knowing and judgment-of-learning different? Evidence from older adults. *Acta Psychologica, 3*(139), 458–64.
Souchay, C., Moulin, C. J. A., Clarys, D., Taconnat, L., & Isingrini, M. (2007). Diminished episodic memory awareness in older adults: evidence from feeling-of-knowing and recollection. *Consciousness and Cognition*, *16*, 769–784.
Swerts, M., & Krahmer, E. (2005). Audiovisual prosody and feeling of knowing. *Journal of Memory and Language*, *53*, 81–94.
Tekcan, A. İ., & Akturk, M. (2001). Are you sure you forgot? Feeling of knowing in directed forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 1487–1490.
Tekcan, A. İ., Topcuoglu, V., & Kaya, B. (2007). Memory and metamemory for semantic information in obsessive–compulsive disorder. *Behaviour Research and Therapy*, *45*, 2164–2172.
Thomas, A. K., Bulevich, J. B., & Dubois, S. J. (2011). Context affects feeling-of- knowing judgments in younger and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 96–108.
Thomas, A. K., Bulevich, J. B., & Dubois, S. J. (2012). An analysis of the determinants of the feeling-of-knowing. *Consciousness and Cognition*, *21*, 1681–1694.
Tuna, S., Tekcan, A. İ., & Topcuoglu, V. (2005). Memory and metamemory in obsessive– compulsive disorder. *Behaviour Research and Therapy*, *43*, 15–27.
Watier, N. N., & Collin, C. A. (2011). Metamemory for faces, names, and common nouns. *Acta Psychologica*, *138*, 143–154.
Watier, N., & Collin, C. (2012). The effects of distinctiveness on memory and metamemory for face-name associations. *Memory*, *21*, 73–88.
Widner, R. L., Smith, S. M., & Graziano, W. G. (1996). The effects of demand characteristics on the reporting of tip-of-the-tongue and feeling-of-knowing states. *American Journal of Psychology*, *109*, 525–538.
Wilkinson, T. S., & Nelson, T. O. (1984). FACTRETRIEVAL2: a Pascal program for assessing someone's recall of general- information facts, confidence about recall correctness, feeling- of-knowing judgments for nonrecalled facts, and recognition of nonrecalled facts. *Behavior Research Methods, Instruments, and Computers*, *16*, 486–488.