City Research Online

City St George's, University of London

Citation: Williams, H. L., Conway, M. A. & Moulin, C. J. (2013). Remembering and knowing: using another's subjective report to make inferences about memory strength and subjective experience. Consciousness and Cognition, 22(2), pp. 572-588. doi: 10.1016/j.concog.2013.03.009

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/15352/

Link to published version: https://doi.org/10.1016/j.concog.2013.03.009

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to City Research Online policy.

City Research Online: http://openaccess.city.ac.uk/ publications@citystgeorges.ac.uk

Remembering and Knowing: Using Another's Subjective Report to Make Inferences About Memory Strength and Subjective Experience

Helen L. Williams[a1], Martin A. Conway[a2], & Chris J.A. Moulin[a3]

[a] Affiliation where the work was carried out: Leeds Memory Group, Institute of Psychological Sciences, University of Leeds, Leeds, LS2 9JT, UK.

[1] Corresponding author. Present address: Department of Psychology, P. O. Box 3050, University of Victoria, Victoria, British Columbia, Canada, V8W 3P5.

Email: helenwilliams098@gmail.com. Phone: +1 250 721 7554. Fax: +1 250 721 8929.

[2] Present address: Department of Psychology, City University, Northampton Square, London, EC1V 0HB, UK. Email: martinconway1@me.com. Telephone: +44 (0)20 7040 8500.

[3] Present address: LEAD - CNRS UMR 5022, Université de Bourgogne, Pôle AAFE, Esplanade Erasme, BP 26513, 21065 Dijon CEDEX.

Email: chris.moulin@me.com. Telephone: +33 (0)3 80 39 39 68.

Abstract

The Remember-Know paradigm is commonly used to examine experiential states during recognition. In this paradigm, whether a Know response is defined as a high-confidence state of certainty or a low-confidence state based on familiarity varies across researchers, and differences in definitions and instructions have been shown to influence participants' responding. Using a novel approach, in three internet-based questionnaires participants were placed in the role of 'memory expert' and classified others' justifications of recognition decisions. Results demonstrated that participants reliably differentiated between others' memory experiences – both in terms of confidence and other inherent differences in the justifications. Furthermore, under certain conditions, manipulations of confidence were found to shift how items were assigned to subjective experience categories (Remember, Know, Familiar, Guess). Findings are discussed in relation to the relationship between subjective experience and confidence, and the separation of Know and Familiar response categories within the Remember-Know paradigm.

Keywords

Remember-Know, subjective experience, autonoetic consciousness, recollection, dual-process, recognition memory

1 Introduction

When the Remember-Know (RK) paradigm was first introduced by Tulving (1985), Remember and Know responses were posited as reflecting autonoetic and noetic consciousness based on retrieval from episodic and semantic memory respectively. In recognition memory tests, old items were categorised as Remembered when the participant retrieved from memory something they had thought or experienced at the time of encoding and categorised as Known when the participant was aware that the item had been on the study list but could not recall anything experienced for the item at that time (Gardiner & Richardson-Klavehn, 2000; Tulving, 1985). Since Tulving's first examination of Remember and Know subjective experiences two main issues have dominated the RK literature: The relationship between subjective experience and confidence, and the relationship between the subjective states of Remembering and Knowing and the underlying processes of recollection and familiarity. The experiments presented here examined both these issues using a novel methodology.

1.1 Subjective experience and confidence

The relationship between subjective experience and confidence has been acknowledged from the conception of the RK paradigm when Tulving (1985) demonstrated that Remembered items were given higher confidence ratings than were Known items. This has been replicated in numerous experiments (e.g., Brewer & Sampaio, 2006; Gardiner & Java, 1990; Rajaram, 1993; Rajaram, Hamilton, & Bolton, 2002; Rotello & Zeng, 2008; Tulving, 1985; and Yonelinas, 2001) with the suggestion that information from subjective awareness is used to judge how confident one is (Gardiner, 2001). The relationship between subjective experience and confidence has also been the subject of debate between advocates of dual-process and single-process models of recognition. Dual-process models (e.g., Tulving, 1985; Mandler, 1980; Yonelinas, 1994, 2002) assume that two

distinct processes underlie recognition and that successful recognition is determined by contributions from both processes. Conversely, single-process accounts propose that recognition relies on only one continuous dimension of familiarity, confidence, or memory strength, and successful recognition is determined by the strength of this single dimension (e.g., Donaldson, 1996; Dunn, 2004). For comprehensive reviews of the literature that summarize the opposing viewpoints see Yonelinas (2002), Diana, Reder, Arndt, and Park (2006), Dunn (2004, 2008), Parks and Yonelinas (2007), and Wixted and Stretch (2004). However, despite continued theoretical debate, how the layperson understands, conceptualises, and acts upon subjective experiences of memory and confidence in memory are still not well understood.

Experiments that have compared RK and confidence judgments have shown different patterns of responding elicited by these two judgment types. Early studies operationalised confidence as a two-category scale of Sure-Unsure and compared proportion of items assigned against proportion assigned to Remember and Know. Using this comparison, these two judgment types were shown to not elicit the same patterns of responses using word/non-word (Gardiner & Java, 1990; Rajaram et al., 2002) or masked priming manipulations (Rajaram, 1993; and with the addition of a Guess category Tunney & Fernie, 2007). More recently, differences between RK and confidence judgments have been demonstrated by comparison of verbal reports at recognition to 'think-aloud' verbalisations made during study (McCabe, Geraci, Boman, Sensenig, & Rhodes, 2011). Here justifications for Remembered items were more likely to contain recollection of details verbalised during encoding than were justifications for high-confidence items . Taken together, these findings demonstrate that while judgments of confidence and subjective experience may be interrelated, RK judgments are not made solely on the basis of confidence and the two are not "experimentally interchangeable" (Rajaram et al., 2002, p. 234).

RK and confidence judgments have also been compared using a larger scale to measure confidence and analysis has focused on the resulting receiver-operating-characteristic (ROC) curves. Using this methodology, higher confidence has been consistently found to be associated with Remember responses compared to Know responses (Rotello, Macmillan, & Reeder, 2004; Rotello, Macmillan, Reeder, & Wong, 2005; Slotnick, 2010; Wixted & Stretch, 2004; Yonelinas, 2001; Yonelinas, Dobbins, Szymanski, Dhaliwal, & King, 1996). Different modelling approaches have aimed to determine whether the data are better explained by single- or dual-process models (e.g., Donaldson, 1996; Dunn, 2004, 2008; Gardiner, Ramponi, & Richardson-Klavehn, 2002; Macmillan, Rotello, Verde, 2005; Wixted & Mickes, 2010). In a review, Gardiner (2008) concluded that most of these available models provide a reasonably good fit to the data but that as the technical complexity of the models has increased it has become more and more difficult to see how to discriminate between them empirically (see also O'Connor, Guhl, Cox, & Dobbins, 2011). Rather than attempting to provide evidence supporting one or other of the models, the experiments presented here examine how people make and understand judgments of subjective experience and confidence. In particular, we were interested in whether the content and nature of people's subjective reports was enough to 'recover' information about confidence and experiential state from those who actually performed the memory test. Our goal was to examine how people interpret evidence from experiential reports in terms of both confidence and categories of subjective experience.

1.2 Knowing and familiarity

A second issue of debate within the RK paradigm is the relationship between the states of Remembering and Knowing and the processes of recollection and familiarity. In their review, Gardiner and Richardson-Klavehn (2000) identified interpretation of Know responses as "the most

vexatious problem in the remember/know paradigm" (p. 238). The root of this problem lies in, firstly, whether Know responses are defined to participants in terms of familiarity or certainty; and secondly, whether Know responses are interpreted as reflecting an underlying process of familiarity or a state of knowing. For example, some researchers choose to ask participants to make Remember-Familiar judgments instead of Remember-Know. Donaldson, MacKenzie, and Underhill (1996) assert "…*familiar* rather than *know* was used to indicate nonrecollection, because the word *know* carries a connotation of certainty that is inconsistent with a confidence rating that indicates lack of certainty. Participants find it hard to say that they are unsure that an item was there but that they know it was" (p. 487, italics in original). Other researchers choose to encompass both familiarity and knowing within one response category, for example, Kelley and Jacoby (1998) define Knowing as "…the inability to recollect any details of the study presentation in combination with *a feeling of familiarity or certainty* that the word was studied" (p. 134, italics added). The issue of how Know responses are interpreted in terms of a familiarity process or a subjective state of knowing is not helped by many research reports not including the exact wording used to define response categories to participants.

Some researchers have separated Know and Familiar as response options. In their study of student learning, Conway, Gardiner, Perfect, Anderson, and Cohen (1997) found that participants could differentiate K and F responses; a K response indicating that they 'just knew' the answer. Students took multiple-choice question exams following four psychology lecture courses and students assigned answers to one of four categories: Remember (R), Know (K), Familiar (F), or Guess (G). At initial testing, higher performing students designated more answers as Remember than did poorer performing students; however at re-test, these students assigned more answers to Know than to Remember. Conway et al. discuss this 'R-to-K shift' as reflecting a change in

knowledge representation from episodic to semantic memory brought about by loss of episodic details from memory and the emergence of conceptual organisation. Evidence of the R-to-K shift has been demonstrated in similar studies of student learning by Barber, Rajaram, and Marsh (2008) and Herbert and Burt (2001, 2003, 2004) and a study exploring learning of rare word definitions by Dewhurst, Conway, and Brandt (2009). Dewhurst et al. also found that Remember and Know responses did not differ in accuracy or confidence, though both were significantly more accurate and associated with higher levels of confidence than Familiar and Guess responses.

It could be argued that the separation of Know and Familiar might only apply to learning of rich materials over a longer time period (e.g., Conway et al., 1997; Herbert & Burt, 2001, 2003, 2004). In typical recognition experiments with lists of unrelated words, where memory is measured shortly after study, it would be difficult for participants to integrate or semanticise the studied information into any body of knowledge and therefore the experiential state of 'Knowing' might not appear to be applicable. However, Dewhurst et al. (2009) asked participants to learn rare word definitions specifically because these were unrelated facts that would be less easy to integrate into a wider schema. While these materials are arguably more meaningful than word lists, they are still less meaningful than the academic material used previously. Dewhurst et al.'s (2009) participants assigned 10% of items to Know at first test – a situation analogous to a single-time-point recognition experiment. While this proportion is lower than the 22% assigned to Know in Conway et al. (1997), it demonstrates that even for less meaningful learning, participants considered the Know response to reflect their experiential state for some items only five minutes

after study with no opportunity for integration into a wider body of knowledge[1]. These 10% of unrelated facts were 'just known' without any recollection or feelings of familiarity regarding the study episode only a short time after said study episode. This suggests that the separation of Know and Familiar might well be applicable to the subjective experiences in typical episodic recognition paradigms where the materials have lower intrinsic meaning (e.g., lists of unrelated words). Indeed, unpublished experimental work by the current authors has demonstrated that the four categories can be used reliably in episodic tasks and that K and F responses differentiate on key characteristics such as accuracy (Williams, Conway, & Moulin, 2013a, 2013b; Williams, Moulin, & Conway, 2009).

Participants thus appear to appreciate the differences between the experiential states of Remember, Know, Familiar, and Guess. Also, for a range of materials, the subjective experiences associated with learning demonstrate consistent patterns over time. In the current experiments, we examined whether these concepts map onto the evidence provided in justifications of recognition that differ in subjective experience and confidence level. Here our focus was on how individuals assess subjective information relating to memory experiences outside the context of a memory task.

1.3 The influence of terminology and confidence on Remember-Know judgments

A related literature considers how definitions of subjective experience categories can influence participants' RK judgments. Rotello et al. (2005) and McCabe and Geraci (2009)

---

[1] Additionally, in Dewhurst et al. (2009) 78% of correct responses at Time 1 were Remembered so there were only 28% of items that could be assigned to any other category of subjective experience.

compared traditional Remember instructions (after Rajaram, 1993) with more conservative Remember instructions that constrained recollection to the study episode. More conservative instructions led to participants using the Remember response less frequently. Examining Know definitions, Geraci, McCabe and Guillory (2009) emphasised Know responses as either highly confident or as less confident. When confidence was emphasised, Remembering and Knowing were differently influenced by words and non-words whereas confidence judgments were not. When confidence was not emphasised, patterns were similar for Remember-Know and Sure-Unsure responses. Geraci et al. (2009) concluded that the wording of instructions can have important theoretical implications for our understanding of Remembering, Knowing, and confidence. In the current experiments, source-specific instructions were included in definitions of Remember, Know, Familiar, and Guess responses, and Remember, Know, and Familiar definitions were each accompanied by a real-world example of the subjective experience, see Table 1. Geraci et al. (2009) also advocated the use of manipulation checks post-recognition to ensure that participants are utilising the response categories as instructed. One common manipulation check is to ask participants to justify their recognition responses, i.e., state *why* they said they recognised that item. Gardiner, Richardson-Klavehn, and Ramponi (1997) used this form of manipulation check, analysed the content of the obtained statements, and published a list of 270 justifications of Remember, Know, and Guess recognition decisions in Gardiner, Ramponi, and Richardson-Klavehn (1998). These justifications were used as stimuli in the current experiments[2].

In Gardiner et al. (1997), after participants had assigned each item to Remember, Know, or Guess, the experimenter chose two responses from each category at random and asked the

[2] We are grateful for John Gardiner's permission to use these materials.

participant to explain their response. Importantly, the emphasis was on what led to the recognition decision; participants were not asked to justify why they had assigned a word to a particular category of subjective experience. Two expert raters then classified the justifications by their salient characteristics (Gardiner et al., 1998). Remember justifications were found to involve: intra-list associations, extra-list associations, item-specific images, and the item's physical features; which Gardiner et al. related to participants' use of encoding strategies. Remember justifications also included self-reference; this was suggested as reflecting items automatically triggering awareness of a personal memory. For Know and Guess justifications, transcripts were shorter and used rather limited vocabulary. Know justifications lacked recollection of specific contextual details and were instead characterised by feelings of familiarity, just knowing, thinking a word occurred, or reporting of the absence of recollective details. Guess justifications demonstrated inferences, speculations, and other judgmental strategies employed by participants.

To assess the relationship between subjective experience and confidence expressed in the justifications, Gardiner et al. (1998) coded which Know and Guess justifications indicated certainty or uncertainty by counting phrases such as 'sure'/'not sure', 'confident'/'not confident', or 'I know'/'I think'. For Know responses, 25% of justifications were rated as indicating certainty, and 11-20% as indicating uncertainty. For Guess justifications, 72-77% were rated as uncertain and none were rated as certain. Through their exploration of how participants justified recognition decisions Gardiner et al. demonstrated that justifications for different subjective experiences reflected access to different memory processes at study and retrieval. Furthermore, the justifications also revealed differences in confidence associated with different subjective experiences.

1.4 Aims of the current experiments

The current experiments explore how people understand reports of memory experiences in terms of confidence and categories of subjective experience. The novel approach here is that instead of asking participants about their own 'first-person' experiences during a memory task, we placed them in the role of 'memory expert' and asked them to classify other people's reports of recognition decisions. This 'third-person' approach is adopted to examine whether categories of subjective experience can be thought of as natural conceptualisations of memory experiences that people readily use and apply. Here three internet-based questionnaires are reported, in which the justification statements of Gardiner et al. (1997, 1998) were used as stimuli. In Experiment 1, participants were asked to assign a confidence rating to each justification statement. These confidence ratings were then used to select justifications for use as stimuli in Experiments 2 and 3 where participants were asked to categorise justifications as Remember, Know, Familiar, or Guess when confidence was manipulated. Our goal was to examine whether, by changing the confidence associated with an item, it was possible to shift interpretations of the category of subjective experience.

In experimental tasks, Remember responses are typically associated with higher confidence than Know responses (Brewer & Sampaio, 2006; Gardiner & Java, 1990; Rajaram, 1993; Rajaram et al., 2002; Rotello & Zeng, 2008; Tulving, 1985; and Yonelinas, 2001). Experiment 1 was designed to examine whether this relationship was replicated when it is not the participants' own contents of memory under study – whether 'third-person' understanding of subjective experiences links experiential state to confidence in the same manner that 'first-person' tasks have found. If people are able to interpret others' justifications in terms of the underlying memory processes, and

these are linked to confidence, then each category of justification statement should be assigned different confidence ratings.

Experiments 2 and 3 manipulated the confidence associated with a report of subjective experience in a method somewhat akin to the experimental paradigms of Geraci et al. (2009), McCabe and Geraci (2009), and Rotello et al. (2005). In Experiment 2, confidence was manipulated through use of justification statements that had obtained high, medium, and low confidence ratings in Experiment 1. In Experiment 3, confidence was manipulated by each justification statement being accompanied by a confidence value. In both experiments, instead of giving a confidence rating for the item, participants assigned justifications to a subjective experience category. If participants' interpretations of the memory processes underlying the justifications are influenced by the confidence associated with the justification, then this should be reflected in how the items are assigned to the categories of subjective experience.

2 Experiment 1

In this experiment, questionnaire respondents were asked to rate how confident they thought a previous participant had been in their recognition decision based on only their justification statement. If participants' confidence ratings are reliably different for items reflecting different categories of subjective experience it would demonstrate that people are able to understand and interpret others' reports of memory experiences in a systematic way, and indeed the relationship between confidence and RK responses is a genuine one based on the use of subjective thoughts and experiences to generate confidence. We were particularly interested in the separation of Know and Familiar categories of subjective experience (e.g., Conway et al., 1997). If Know and Familiar justifications are seen to reflect different levels of confidence this will add support to the validity of using these two separate response categories in recognition memory

experiments. How Know and Familiar items were operationalised separately in the current experiments is discussed in the Results of Experiment 1 (Section 2.2.1).

2.1 Method

2.1.1 Participants

Data were collected using an online questionnaire, active from April 2008 to September 2009. It was advertised on international experiment websites and emailed to Leeds Memory Group (LMG) participant lists. Some participants were University of Leeds psychology undergraduates and were given participation credit as compensation. Full data sets from 309 participants were obtained (225 female, 84 male; *mean age* = 28.1, *SD* = 10.78, *range* = 16 to 65).

2.1.2 Materials and Design

The cue words and justification statements used as stimuli were those published in Gardiner et al. (1998). For use in the current questionnaire, all 270 items were sorted into 10 lists of 27 items for ease of completion by participants online. Each list contained equal numbers of Remember, Know, and Guess items (9 of each). Participants were randomly assigned to a list when they accessed the questionnaire and within each list items were presented in random order.

2.1.3 Procedure

On accessing the questionnaire, participants were first presented with ethics and consent information. Experiment instructions then stated participants would be presented with statements that were responses made by people in a previous memory experiment where people had to learn a series of words and were later tested using an Old-New recognition paradigm. In this paradigm, for each word recognised as Old, the person had then been asked to justify their response – why they thought they recognised that word. Participants in the current experiment were instructed that they would be shown the justification statements and that their task was to rate how confident they

thought the person had been about their recognition. An example was then given before participants entered their demographic information and were shown the first item.

For each item, participants saw the cue word and justification statement accompanied by the question, "How confident do you think this participant was that they had accurately recognised this word?" They made their confidence rating by selecting a number from 0 to 100 (in increments of five) from a drop-down box. After all 27 items, participants were given debrief information, were permitted to provide comments about the experiment, and were asked whether they would like to provide their email address so they could be emailed about future online experiments.

2.2 Results and Discussion

To examine whether confidence levels associated with justification statements differed the mean confidence rating for each of the 270 items was calculated. The mean confidence ratings made to items from each subjective experience category (from the original Gardiner et al., 1998, paper) were then examined. Remember items received the highest confidence ratings (M = 81.62%, SD = 7.21), followed by Know (M = 52.56%, SD = 13.50), and then Guess (M = 34.34%, SD = 11.29). ANOVA demonstrated a significant main effect of subjective experience (as a between-items factor) $F(2,267) = 424.38$, $p < .001$, and t-tests showed significant differences between confidence ratings assigned to each category of subjective experience (all $p < .001$). These results suggest that not only is there a relationship between subjective experience and confidence when it the contents of one's own memory that is being assessed but that judgments of confidence based on the experiential reports of others show a reliable relationship also. People understand the relationship between subjective experience and confidence when asked to judge

reports of another's memory experiences in the same way as is observed in experimental paradigms.

2.2.1 Splitting Know into Know and Familiar

The problem of the unity of the concept of Knowing is exemplified by the fact that Gardiner et al. (1998) rated 25% of Know responses as indicating certainty and up to 20% as indicating uncertainty. This suggests a great deal of variance within the subjective experiences underlying Know responses in their experiment. One aim of the current experiments was to examine whether some of the Know items provided by Gardiner et al.'s (1997, 1998) participants perhaps leaned more towards being a justification of Familiarity rather than of Knowing. To address this, two of the current researchers (HLW and CJAM) independently categorised Gardiner et al.'s (1998) original 90 Know justifications as reflecting either a Familiar or Know recognition response based on the definitions of Know and Familiar shown in Table 1. Comparing the independent ratings revealed that categorisation had matched on 74 of the 90 items giving an inter-rater reliability of .82. For the remaining 16 items, each was discussed until consensus was reached. Of the 90 items, final classification of expert ratings was that 47 were justifications reflecting a Know response and 43 were justifications based on Familiarity (see the Supplementary Materials for which were classed as Know and which as Familiar). These separated Know and Familiar items were examined independently in all the experiments presented here.

[Insert Table 1 about here]

In Experiment 1, to examine whether the split of the Know items provided by Gardiner et al.'s (1998) participants into separate Know and Familiar categories was reflected in the confidence values assigned to items, the above analysis was repeated using the four categories of subjective experience. ANOVA again demonstrated a significant main effect of subjective

experience (as a between-items factor), $F(3,266) = 321.43$, $p < .001$, and t-tests showed significant differences between confidence ratings assigned to each category of subjective experience (all $p < .001$). Remember items[3] received the highest confidence ratings, followed by Know ($M = 58.18\%$, $SD = 11.29$), then Familiar ($M = 46.42\%$, $SD = 9.47$), then Guess. Different levels of confidence were reliably associated with the separate Know and Familiar categories of subjective experience. This supports the experimental findings of Brewer and Sampaio (2006), Dewhurst et al. (2009), and Kihlstrom, Kim, and Dabady (1996) that Know responses are associated with higher levels of confidence than Familiar responses. The current findings suggest that people are also able to differentiate these two experiential states based on others' memory reports and, at least in part, this discrimination is related to the level of confidence associated with the subjective experience. These findings also relate to Gardiner et al.'s (1998) original expert ratings, which concluded that 25% of Know justifications were rated as indicating certainty but ratings of how many indicated uncertainty varied between 11 and 20%. Separation of the original Know items into Know and Familiar has revealed differences in confidence associated with justifications that were all originally provided to Know recognition responses. This adds further validity to these two subjective states being provided as separate response options in experimental paradigms.

This experiment has demonstrated that different categories of subjective experience justifications receive different ratings of confidence. Although this experiment did not take a first-person approach to understanding subjective experience as it asked participants to rate others'

---

[3] The mean confidence ratings made to Remember and Guess items were identical to those reported in the previous analysis as these justifications types were not affected by the splitting of Know into Know and Familiar.

memory descriptions, the findings do lend support to the idea that it is subjective awareness that gives rise to confidence (Gardiner, 2001; Tulving, 1985). However, this experiment only examined the relationship between subjective experience and confidence in one direction; Experiments 2 and 3 explore whether different levels of confidence can be categorised as different types of subjective experience.

3 Experiment 2

As discussed earlier (Section 1.3), manipulations of source constraint in definitions of Remembering or emphasis on confidence in definitions of Knowing have been shown to result in altered patterns of RK responses (Geraci et al., 2009; McCabe & Geraci, 2009; Rotello et al., 2005). Here we were interested in whether similar manipulations of confidence influence how participants interpreted other people's justification statements in terms of subjective experience categories. To manipulate confidence, the confidence ratings obtained in Experiment 1 were used to select justifications associated with high, medium, and low levels of confidence. The task for participants was to assign justifications to Remember (R), Know (K), Familiar (F), and Guess (G) categories.

If the confidence associated with a justification influences how participants interpret the subjective experience expressed in that statement, then whether the item is assigned to R, K, F, or G should differ across confidence levels. For example, a statement that originally justified a Know response (in Gardiner et al., 1998) but was perceived as expressing low confidence (in Experiment 1) may not be consistently assigned to the Know category. In addition, the influence of confidence may vary according to item type. For example, the evocative recollections, associations, item information, and self-reference associated with Remember justifications may mean that categorisation of these items is unaffected by confidence and they will be consistently

assigned to R. However, as Know, Familiar, and Guess justifications lack the detail of Remember justifications, how they are assigned to subjective experience may be more greatly influenced by the confidence associated with the justification.

More broadly, if the confidence associated with a justification influences how that justification is interpreted in terms of subjective experience then this would suggest that confidence plays an important role in how people understand memory experiences. On the other hand, if confidence does not influence how people categorise others' memory justifications, and the original subjective experience of the item can be recovered, this would suggest that confidence does not play an important role in how people understand and interpret the phenomenology of memory.

3.1 Method

3.1.1 Participants

Data were collected using an online questionnaire, active from January to September 2009. Advertisement was undertaken in the same manner as in Experiment 1. Full data sets from 502 participants were obtained (388 female, 111 male, 3 no response; *mean age* = 26.71, *SD* = 12.40, *range* = 16 to 85).

3.1.2 Materials and Design

The cue words and justification statements used as stimuli were a selection of those used in Experiment 1. Using the mean confidence ratings obtained from participants in Experiment 1, within each subjective experience category (Remember, Know, Familiar, and Guess), statements were sorted from highest to lowest confidence assigned. The two statements with the highest,

medium[4], and lowest confidence values were selected from each category giving a set of 24 items. A second selection of 24 items was collated from items with the next lowest, medium, and highest confidence ratings to be used as a second version of the questionnaire. The two versions were then swapped online intermittently in order that approximately equal numbers of participants completed each. On each version, items were presented in random order for each participant. Full data sets were obtained from 248 participants for Version 1 and from 254 participants for Version 2. As no differences were observed between the data from the two versions all analyses were conducted on both versions together. The statements used in this experiment are marked in the Supplementary Materials and the maximum, minimum, and mean confidence ratings for High, Medium, and Low confidence items for each subjective experience category are shown in Table 2.

[Insert Table 2 about here]

As shown in Table 2, confidence ratings did overlap across subjective experience categories. For example, Remember justifications that had achieved the Lowest confidence ratings (M = 58.28) had been assigned similar confidence ratings to both Medium confidence Know items (M = 56.40) and High confidence Guess items (M = 57.66). This is interesting in itself, as whilst these justifications were rated as reflecting the same levels of confidence in recognition in Experiment 1, they had originally justified different subjective experience responses in Gardiner et al.'s (1997, 1998) experiment.

3.1.3 Procedure

The procedure for this experiment closely followed that of Experiment 1 though instead of making confidence ratings, here participants were presented with the cue word and justification

[4] Based on the median.

statement and asked to classify that item as Remember, Know, Familiar, or Guess. Initial instructions regarding how the justifications statements were obtained matched those given for Experiment 1. Participants were instructed that their task was to make a judgment about the person's recognition decision by classifying their statement as Remember, Know, Familiar, or Guess. Full definitions of the categories were provided (see Table 1) and it was emphasised that participants should ensure that they fully understand the definitions before continuing as the definitions would not be shown again. It was suggested that participants could write down the definitions to refer to later. Participants were also told that they may only select one category for each statement but that they should try to use all the categories at least once, and to read all the items carefully as some of the justifications would be easier to categorise than others. Participants then entered their demographic information before being shown the first item. For each item, participants were shown the cue word and justification statement accompanied by the question "Which recognition category?" and the options Remember, Know, Familiar, and Guess; response options were always presented in this order. Participants made their decisions by selecting the appropriate radio button. After all 24 items participants were debriefed as in Experiment 1.

3.2 Results and Discussion

This experiment examined whether participants were able to reliably match justification statements to their original category of subjective experience and whether categorisation would differ depending on the confidence level assigned to the statement by participants in Experiment 1. Analysis took the form of items-analysis with item type (REMEMBER, KNOW, FAMILIAR, GUESS) and confidence level (High, Medium, Low) as between-items factors and the subjective experience response of the participant (R, K, F, G) as a within-items factor. For clarity, within the Results section only, when CAPITALISED the terms REMEMBER, KNOW, FAMILIAR,

GUESS refer to the original subjective experience of the item. When only the initial is provided (R, K, F, G) this refers to the subjective experience response of the participant in the current task, i.e., what category of subjective experience they assigned the item to.

The proportion of items within an item type and a particular level of confidence that were assigned to each of the subjective experience response categories were calculated across participants, e.g., the proportion of High confidence REMEMBER items that were assigned to R by participants. Firstly a 4(item type) x 3(confidence level) x 4(subjective experience response) ANOVA was performed to examine interactions between our between-items variables and our within-subjects responses[5]. Separate ANOVAs and further comparisons were performed to

---

[5] In this analysis, proportions of subjective experience responses (R, K, F, and G) summed to 1 within each confidence level and therefore between-items factors of confidence level and item type could not be calculated. Although these proportions are not independent, we employed mixed ANOVAs in our initial examination of data patterns because we were interested in the interactions between our between-items variables and our within-subjects variable; we were not interested in main effects of item type or confidence. This method follows the statistical approach used by other studies including four subjective experience response categories (Conway et al., 1997; Herbert & Burt, 2001, 2003, 2004). Furthermore, this approach is analogous to a Chi Square Test of Independence and is subject to the same constraints. We do not believe that the variation in patterns of responses that we observed across levels of our two between-items conditions is a result of the dependent nature of the response categories. From our patterns of data it seems unlikely that any other statistical approach would result in very different conclusions.

examine patterns within the different levels of confidence and item types. Greenhouse-Geisser values are reported when ANOVA did not meet assumptions of sphericity.

The 4(item type) x 3(confidence level) x 4(subjective experience response) ANOVA demonstrated a significant main effect of subjective experience response, $F(2.03,73.12) = 9.11$, $p < .001$, and significant two-way interactions between item type and subjective experience response, $F(6.09,73.12) = 41.12$, $p < .001$, and confidence level and subjective experience response, $F(4.06,73.12) = 14.13$, $p < .001$. These were qualified by a significant three-way interaction between item type, confidence level, and subjective experience response, $F(12.19,73.12) = 5.55$, $p < .001$.

The means for the significant interaction between item type and subjective experience response are shown in Figure 1. When confidence level is not considered, the majority of REMEMBER and FAMILIAR items were appropriately allocated to their item type. Conversely, for KNOW items nearly as many were assigned to F as to K, and for GUESS items a large proportion were assigned to F instead of G.

[Insert Figure 1 about here]

The significant three-way interaction is explored in the following sections. Firstly, within each item type a 3(confidence level) x 4(subjective experience response) ANOVA was performed. If there was a significant interaction of confidence and subjective experience response for that item type then separate 3(confidence level) ANOVAs were conducted that examined use of each of the four subjective experience responses across confidence levels. Results of these analyses are reported in the next four sections.

3.2.1 Assignment of High, Medium, and Low confidence REMEMBER items to R, K, F, and G

How confidence level interacted with subjective experience response for REMEMBER items is shown in Figure 2a. For statements originally justifying a REMEMBER recognition decision in Gardiner et al.'s (1998) experiment, irrespective of it receiving a High, Medium, or Low confidence rating by participants in Experiment 1, the majority of REMEMBER items were appropriately assigned to R. A 3(confidence level) x 4(subjective experience response) ANOVA demonstrated a significant main effect of subjective experience response, $F(3,27) = 209.50$, $p < .001$, and a significant interaction between confidence level and subjective experience response, $F(3.75,16.86) = 3.64$, $p = .028$.

[Insert Figure 2 about here]

ANOVAs comparing use of each subjective experience response category across confidence levels demonstrated that only for F and G responses was a significant main effect of confidence level found: F: $F(2,9) = 9.69$, $p = .006$, G: $F(2,9) = 6.36$, $p = .019$. The likelihood of an item being assigned to F or G decreased as confidence level increased. A greater proportion of REMEMBER items were assigned to F if they were Medium confidence compared to High confidence, $t(6) = 3.96$, $p = .007$, and Low compared to Medium levels of confidence approached significance, $t(6) = 2.21$, $p = .069$. A greater proportion of REMEMBER items were assigned to G if they were Low confidence compared to Medium confidence, $t(6) = 2.44$, $p = .05$.

In sum, the majority of REMEMBER items were consistently categorised as R, regardless of the confidence associated with the justification. In addition, the proportion of REMEMBER items that were assigned inappropriately to K did not differ with confidence, only the proportion of REMEMBER items inappropriately assigned to F and G differed depending on the confidence level associated with the justification.

The suggestion that Remember items would be appropriately classified as Remember regardless of the confidence level of the justification was supported, as at all levels of confidence the majority of items were assigned to R. We suggest that this is due to the nature of justifications made in support of recognition responses based on recollection. As Gardiner et al. (1998) reported, Remember justifications typically included details such as intra- and extra-list associations, item-related information, and self-reference. These rich, evocative, details – which Gardiner et al. reported that participants were able to bring to mind with ease, are suggested as being critically important for identifying these items as Remember. In examining the justifications employed in this experiment (shown in the Supplementary Materials), even the Remember items that had been given the lowest confidence ratings by participants in Experiment 1 included many such associations and details. For example, the lowest confidence Remember item involved intra-list association: *Ape "When it came up I remember thinking that there were lots of words with three letters".*

3.2.2 Assignment of High, Medium, and Low confidence KNOW items to R, K, F, and G

In contrast to the results for REMEMBER items shown in Figure 2a, the patterns for assignment of KNOW items to subjective experience categories demonstrate that categorisation of KNOW items was strongly influenced by confidence level, see Figure 2b. A 3(confidence level) x 4(subjective experience response) ANOVA demonstrated a significant main effect of subjective experience response, $F(3,27) = 6.79$, $p = .001$; and a significant interaction between confidence level and subjective experience response, $F(6,27) = 10.30$, $p < .001$.

For KNOW items, ANOVAs comparing subjective experience responses across confidence levels demonstrated that a significant main effect of confidence level was evident for K: $F(2,9) = 19.55$, $p = .001$, F: $F(2,9) = 12.06$, $p = .003$, and G: $F(2,9) = 5.58$, $p = .027$, responses, and R

responses approached a significant main effect of confidence, $F(2,9) = 3.45$, $p = .077$. The likelihood of a KNOW item being assigned to K fell with confidence level. A greater proportion of items were assigned to K if they were High confidence compared to Medium confidence, $t(6) = 3.39$, $p = .015$, and if they were Medium confidence compared to Low confidence, $t(6) = 3.82$, $p = .009$. This pattern was reversed for the likelihood of a KNOW item being assigned to F as a greater proportion of Medium confidence items were assigned to F than were High confidence, $t(6) = 7.00$, $p < .001$. The likelihood of a KNOW item being assigned to R was greater for Medium confidence KNOW items than for Low confidence KNOW items, $t(6) = 3.10$, $p = .02$. No other comparisons were significant.

In sum, for KNOW items confidence influenced how participants assigned items to categories of subjective experience. As confidence level increased, the proportion of KNOW items assigned to K increased, as did the proportion assigned to R. In addition, the proportion of KNOW items assigned to F and G fell as confidence level increased.

3.2.3 Assignment of High, Medium, and Low confidence FAMILIAR items to R, K, F, and G

The pattern of categorisation of FAMILIAR items (Figure 2c) parallels that for REMEMBER items (Figure 2a); between .50 and .60 of statements originally justifying a FAMILIAR recognition decision were appropriately assigned to the F category of subjective experience irrespective of whether they were of High, Medium, or Low confidence.

For FAMILIAR items the 3(confidence level) x 4(subjective experience response) ANOVA demonstrated a significant main effect of subjective experience, $F(3,27) = 29.22$, $p < .001$, but no interaction, $F(6,27) = 1.09$, $p = .40$. Collapsed across confidence level, a greater proportion of FAMILIAR items were assigned to F than to any other category of subjective experience, all $p < .001$. No other comparisons were significant.

The differential results obtained for the newly separated KNOW and FAMILIAR items by splitting Gardiner et al.'s (1998) participants' Know items based on expert ratings provide support that these two categories of subjective experience can be considered independent. Assignment of KNOW items to subjective experience categories differed depending on the confidence level associated with the item whereas FAMILIAR items were consistently appropriately categorised as F. This is in line with the findings of McCabe and Geraci (2009), Rotello et al. (2005), and Geraci et al. (2009), which demonstrated that the wording of RK definitions influenced participants' use of the categories. Geraci et al. found different patterns of responding when confidence was or was not emphasised in Know definitions. Though Geraci et al. did not argue for two separate states of Knowing and Familiarity, their results did show that different patterns of subjective experience response were obtained when the definition of Knowing was altered. Considered alongside the present findings, the confidence level associated with Know has been found to influence patterns of responding both when participants are assessing their own memory experiences and when judging the memory experiences of others'.

The items used in this experiment were those that had received the Highest, Medium, and Lowest confidence ratings for their subjective experience category in Experiment 1. However, examination of the items (see the Supplementary Materials) highlighted differences between the types of reasons involved in justifications given High, Medium, and Low confidence ratings. High confidence KNOW items reflected certainty of recognition, e.g., *Bluebell "I am sure about that one, there were a couple of words which were similar and were part of the category flower"*; whereas Medium and Low confidence KNOW items mentioned absence of recollection, e.g., *Ring "I think I remember seeing it, but there was no link or image. I can't remember feeling anything"*; and one Low confidence item mentioned familiarity. Some FAMILIAR items at each level of

confidence mentioned familiarity (though not all items did), however only High and Low confidence FAMILIAR items mentioned absence of recollection, and only Medium and Low items mentioned uncertainty, e.g., *Harbour "It was familiar, but I was confused. I knew it was there but could not be sure"*. As systematic differences across confidence levels were not evident to a greater extent for KNOW items than for FAMILIAR items the content of the items cannot directly explain why confidence influenced classification of KNOW items but not FAMILIAR items.

3.2.4 Assignment of High, Medium, and Low confidence GUESS items to R, K, F, and G

As is shown in Figure 2d, how GUESS items were assigned to subjective experience category again differed according to confidence level. A 3(confidence level) x 4(subjective experience response) ANOVA on GUESS items demonstrated a significant main effect of subjective experience response, $F(1.38,12.42) = 20.30$, $p < .001$, and a significant interaction between confidence level and subjective experience response, $F(1.38,12.42) = 12.45$, $p = .001$.

ANOVAs comparing subjective experience response categories across confidence levels demonstrated that a significant main effect of confidence level was evident for R, $F(2,9) = 9.89$, $p = .005$, F, $F(2,9) = 8.83$, $p = .008$, and G, $F(2,9) = 16.69$, $p = .001$, subjective experience responses; and K responses approached significance, $F(2,9) = 3.26$, $p = .086$. The patterns in Figure 2d show that the likelihood of a GUESS item being assigned to G increased as confidence level fell. However, only the difference between proportion assigned to Low and Medium confidence levels was significant, with a greater proportion of GUESS items being assigned to G if they were Low confidence compared to Medium confidence, $t(6) = 2.97$, $p = .025$. This pattern was reversed for the likelihood of a GUESS item being inappropriately assigned to F. A greater proportion of Medium confidence GUESS items were assigned to F than were Low confidence GUESS items, $t(6) = 3.09$, $p = .021$. In addition, the likelihood of a GUESS item being

inappropriately assigned to R was higher if it was High confidence as opposed to Medium confidence, $t(6) = 2.79$, $p = .032$. No other comparisons were significant.

In sum, as with KNOW items, for GUESS items the confidence associated with the items strongly influenced how participants assigned these items to categories of subjective experience. As confidence level fell, the proportion of GUESS items assigned to F decreased and the proportion assigned to G increased. Examination of the GUESS items used in this experiment revealed that whereas Low confidence justifications were typically very short and explicitly discussed guessing, uncertainty, or judgmental strategies, e.g., *Father "I kept saying 'no' so I just guessed it was there because you said that 50% of the words were there"*; Medium and High confidence justifications also often mentioned confusion over the certainty or familiarity of a word, e.g., *Flea "I am almost certain that it was there. But not entirely"*. It is suggested that these differences led participants in Experiment 1 to give these items different confidence ratings and participants in Experiment 2 to categorise a large proportion of High and Medium confidence GUESS items inappropriately as F. That some Guess responses reflect strategic decisions based on appropriate response rates while other guesses are based on lower levels of familiarity, or confusion of familiarity (Gardiner et al., 1998), appears to have led participants to categorise GUESS items inconsistently in the current experiment.

This experiment explored individual's understanding of the relationships between how one justifies a memory and the confidence associated with that justification. For REMEMBER or FAMILIAR items, irrespective of these items receiving a High, Medium, or Low confidence rating by participants in Experiment 1, the majority of these items were appropriately assigned to R or F categories respectively. Conversely, the confidence levels assigned by participants in Experiment 1 strongly influenced how KNOW and GUESS items were assigned to categories of subjective

experience. KNOW items appropriately assigned to K increased as confidence level increased and, conversely, GUESS items appropriately assigned to G increased as confidence level fell.

One explanation for the finding that classification of KNOW and GUESS items differed depending on confidence level whereas REMEMBER and FAMILIAR items were less sensitive to manipulated confidence is that REMEMBER and FAMILIAR justifications more directly describe the processes underlying their recognition decisions – recollection and familiarity, and this was evident to participants when they were categorising the statements. Although this could be taken as inferring that separate KNOW and GUESS categories of subjective experience do not accurately assess the processes underlying recognition, an alternative suggestion is that K responses should be conceptualised as subjective experiences reflecting high confidence without recollection, and G responses should be conceptualised as low-confidence familiarity-based responses. As demonstrated by Geraci et al. (2009), different patterns of responding are observed if confidence is emphasised in how Knowing is defined. Additionally, as discussed above, some GUESS items demonstrate confusion concerning the certainty or the source of familiarity of a previously seen word. If Know and Guess experiences are related to a participant's confidence in their evaluation of what is in memory then it follows that these responses would be influenced to a greater extent by confidence level when it is manipulated experimentally.

Overall, these findings demonstrate that when understanding the memory experiences of others, participants' judgments regarding subjective experience were influenced by confidence, and this influence was greater for Know and Guess categories of subjective experience. This novel method of using participants' justification statements as stimuli and exploring the relationship between subjective experience and confidence in a 'third-person' approach has added to the experimental 'first-person' recognition findings of Rotello et al. (2005), McCabe and Geraci

(2009), and Geraci et al. (2009), supporting the assertion that subjective experience judgments are influenced by manipulations of confidence. To further link this novel methodology to the prior experimental findings the confidence associated with each justification statement was manipulated more overtly in Experiment 3.

4 Experiment 3

The aim of this experiment was to further explore the relationship between people's understanding of subjective experience and confidence by providing participants with both a justification statement and a confidence value. However, in this experiment, confidence was manipulated systematically in order to assess whether it influenced the subjective experience category the item was assigned to. Prototypical Remember, Know, Familiar, and Guess justification statements were selected from around each category's mean confidence justification and were presented to participants accompanied by a confidence value that was either appropriate (to that subjective experience category) or inappropriate (from the confidence values of a different subjective experience category). As in Experiment 2, the task for participants was to assign items to Remember, Know, Familiar, and Guess categories of subjective experience.

The focus of this experiment was whether participants based their categorisation decisions on both the justification and the confidence value. In Experiment 2, the confidence associated with a justification statement could not be manipulated while the justification statement itself was held constant; however, this could be done in Experiment 3 – here the statement remained constant while the confidence associated with it differed. If confidence does not influence interpretations of subjective experience, this would result in uniform patterns within item type. On the other hand, high or low confidence values associated with a justification may lead participants to interpret that statement as reflecting a different subjective experience from that which it was

originally justifying. This pattern of results would suggest that confidence plays an important role in how people understand and interpret the phenomenology of memory.

In line with the findings of Experiment 2, it was hypothesised that classification of Know and Guess justifications would be more heavily influenced by the confidence level accompanying them than would Remember and Familiar justifications.

4.1 Method

4.1.1 Participants

Data were collected using an online questionnaire active from October 2009 to February 2010. Advertisement was undertaken in the same manner as in Experiments 1 and 2. Full data sets from 258 participants were obtained (164 female, 92 male, 2 no response; *mean age* = 31.04, *SD* = 17.17, *range* = 16 to 79).

4.1.2 Materials and Design

The cue words and justification statements used as stimuli were a selection from those published by Gardiner et al. (1998). Using the confidence ratings obtained from participants in Experiment 1, within each subjective experience category (Remember, Know, Familiar, and Guess) eight statements from around the mean confidence value were selected as prototypical items for each category (shown in the Supplementary Materials).

As justifications were to be paired with confidence values, eight confidence values were selected around the mean for each subjective experience category. As shown in Table 3, the

ranges of values for each category were non-overlapping and the mean of the selected values matched the original mean confidence as closely as possible[6].

A Latin-square design was used to pair items to confidence values. Within each of the four categories of subjective experience, two justifications were paired with plausible confidence values and two justifications were paired with confidence values derived from the confidence ranges for each of the other original subjective experience categories. For example, of the eight Remember justifications two were paired with Very High confidence values, two were paired with High confidence values, two with Medium confidence values, and two with Low confidence values. This pairing was then repeated for Know, Familiar, and Guess justifications. Using this design, each participant saw 32 items. Four versions of the pairings were created. Careful matching ensured that each confidence value served as a plausible confidence once, for example a Very High confidence of 85 matched to a Remember justification, but on the other versions of the questionnaire that confidence value of 85 was paired with a either a Know, Familiar, or Guess justification. Each confidence value was used only once within each version of the questionnaire. Participants were randomly assigned to a version on accessing the questionnaire and for each version items were presented in random order.

[Insert Table 3 about here]

[6] Means shown in Table 2 do not exactly match those discussed in Experiment 1 as selection of items for this experiment was performed prior to final analysis of Experiment 1. Differences in means varies from 0.30% to 1.30%.

4.1.3 Procedure

The procedure for this experiment closely followed that of Experiment 2 though here participants were presented with both a justification statement and a confidence value on which to base their classification of the item. Initial information regarding how the justification statements were collected from previous participants matched that in Experiment 2. Current participants were instructed that they would be shown justification statements and confidence ratings made by participants in a memory test and that their task was to make a judgment about the person's recognition decision by classifying it as Remember, Know, Familiar, or Guess. Full definitions of the categories were provided (see Table 1) and further instructions matched those in Experiment 2 except participants were also reminded to pay attention to both the justification and the confidence rating for each item. For each item, participants were shown the cue word, justification statement, and confidence rating, accompanied by the question "Which recognition category?" and the options Remember, Know, Familiar, and Guess. They made their decisions by selecting the appropriate radio button. After seeing all 32 items participants were debriefed as in Experiment 2.

4.2 Results and Discussion

Experiment 3 examined whether participants were able to reliably match items to their original category of subjective experience when provided with a confidence value alongside the justification. Analysis again took the form of items-analysis with the item type (REMEMBER, KNOW, FAMILIAR, GUESS) and the manipulated confidence level of the item (Very High, High, Medium, Low) as between-items factors and the response of the participant – the category of subjective experience they assigned the item to (R, K, F, or G), as a within-items factor. The dependent variable was the proportion of items within an item type and a particular level of confidence that were assigned to each of the subjective experience response categories.

The 4(item type) x 4(manipulated confidence level) x 4(subjective experience response) ANOVA demonstrated a significant main effect of subjective experience response, F(2.02,226.42) = 35.22, p < .001, and a significant interaction between item type and subjective experience response, $F(6.09,226.42) = 102.45$, $p < .001$; this interaction is discussed in more detail below. The interaction between manipulated confidence level and subjective experience response was not significant and neither was the three-way interaction between item type, manipulated confidence level, and subjective experience response, both Fs < 1. This non-significant three-way interaction is shown in Figure 3.

[Insert Figure 3 about here]

Figure 3 shows that patterns of responding were stable across confidence level in this experiment. The confidence value that was provided alongside the justification statement did not influence how participants assigned the item to a category of subjective experience; participants primarily based their assignment of the item to R, K, F, or G on the experiential state reported in the justification.

The two-way interaction between item type and subjective experience response was significant however; see Figure 4. The patterns shown in this figure are very similar to those shown for this interaction in Experiment 2 (shown in Figure 1). When confidence level is not considered, the majority of REMEMBER and FAMILIAR items were appropriately allocated to their original subjective experience category. For KNOW items nearly as many were assigned to F as to K, and for GUESS items nearly as many were assigned to F as to G.However, overall the proportion of items categorised as R, K, F, or G were highest when the item was originally a REMEMBER, KNOW, FAMILIAR, or GUESS item respectively, i.e., of the items assigned to each response option the majority come from the appropriate original subjective experience

category. Whilst some KNOW and GUESS items were allocated to F instead of to their appropriate category, the largest proportion of items that were assigned to K and G were KNOW and GUESS items respectively.

[Insert Figure 4 about here]

Examination of the Know items used in this experiment revealed that they typically referred to lack of recollection or uncertainty about recognition; for example, *Tangerine "I recognised it as a word from yesterday, but I cannot really remember what I thought, I could not remember seeing it on the screen but I was sure it was there yesterday".* None of the Know items included the word 'familiarity' so it was not the use of this word that led to Know items being assigned to Familiar. Instead it is suggested that the uncertainty evident in some of the Know items (see *Tangerine* above) is what led participants to categorise some items as Familiar instead of Know. This could suggest that the separation of the traditional Know category into separate Know and Familiar categories could require more refinement of definitions; however, evidence from previous research that has separated these categories (e.g., Brewer & Sampaio, 2006; Conway et al., 1997; Dewhurst et al., 2009; Herbert & Burt, 2001, 2003, 2004; Kihlstrom et al., 1996) suggests that many aspects of this split are reliable. We return to this issue in the General Discussion.

All Guess items used in this experiment referred to uncertainty about recognition; for example, *Harbour "I lived by the sea all my life, so I was not sure whether I have encountered that word here or whether it is to do with home".* Only one Guess item included the word 'familiar'. Furthermore, while Gardiner et al. (1998) demonstrated that Guess justifications can reflect inferences or judgmental strategies, only one item in the current experiment contained an inference concerning the study list: *Harp "It seemed that there were quite a few musical*

instruments, so I took a guess that it came up". Instead of judgmental strategies and inferences, it is suggested that the uncertainty about why an item felt familiar inherent in the Guess items in this experiment is what led many of these items to be categorised as Familiar.

In sum, in this experiment participants were provided with a confidence value alongside the justification statement that they were to assign to a category of subjective experience. The confidence values did not influence participants' assignment of items; there was no interaction between item type, manipulated confidence level, and subjective experience response, and no interaction between manipulated confidence level and subjective experience response. Participants focused their categorisation on the experiential phenomenology in the statements and the majority of classifications were appropriate. Furthermore, the different patterns of categorisation of Know and Familiar items demonstrate that there is something within these justification statements that enables people to categorise them differently; this issue is returned to in Section 5.2.

5 General discussion

We examined two central issues in the RK literature in these experiments using the novel methodology of asking participants to make judgments about the subjective experiences in others' memory reports. Results speak to both the relationship between subjective experience and confidence, and how separated Know and Familiar states of awareness map onto the underlying memory processes of recollection and familiarity.

5.1 Subjective experience and confidence

Experiment 1 demonstrated that mean confidence assigned to statements originally made to justify Remember, Know, Familiar and Guess responses differed significantly. Remember items were assigned higher confidence ratings than Know items, Know were higher than Familiar,

Familiar higher than Guess. Thus people appear to understand the relationship between subjective experience and confidence from others' memory experiences in the same way as is observed in experimental paradigms (Gardiner & Java, 1990; Dewhurst et al., 2009; Rajaram, 1993; Rajaram et al., 2002; Rotello & Zeng, 2008; Tulving, 1985; and Yonelinas, 2001). Extending this, Experiment 2 demonstrated that when the confidence associated with an item is manipulated intrinsically to the justification (High, Medium, and Low confidence statements selected as stimuli) this influenced how items were assigned to categories. Remember and Familiar items were little influenced by associated confidence level, whereas Know and Guess items were often categorised according to confidence. This suggests that the qualitative nature of Remember and Familiar justification statements are not sensitive to manipulations of confidence while Know and Guess are.

However, when an explicit confidence rating was provided alongside a justification statement in Experiment 3 it had no influence on the categorisation of items. At all levels of confidence, Remember and Familiar items were again consistently assigned to Remember and Familiar, while whatever confidence value was presented alongside Know and Guess justification statements a large proportion of items were assigned to Familiar, paralleling the results obtained for Medium confidence items in Experiment 2.

Considered together, when confidence is manipulated as an internal attribute (Experiment 2) it influences how the item is assigned to subjective experience category, especially for Know and Guess items. However, when confidence is manipulated as a linked but external feature (Experiment 3) it does not influence how items are assigned. This is taken as support for the view that confidence derives from subjective experience as proposed by Tulving (1985) and Gardiner "…it is surely the subjective state of awareness that gives rise to confidence in memory, not

confidence that gives rise to the state of awareness" (Gardiner, 2001, p. 1356). When both a statement and a confidence value were provided, categorisation appeared to be largely based on the justification report rather than the confidence value. It is suggested that this occurred because participants used their natural understanding of memory experiences to determine that subjective experience is of more importance than confidence.

Given the current findings, it seems difficult to suggest that confidence is merely the driving force behind judgments of subjective experience (cf. Dunn, 2008). Even when people are faced with categorising recognition decisions on the basis of third-party information, they are more likely to use qualitative information about the retrieval than the easily-interpreted confidence value. Instead, we suggest that people habitually use the experiential state, and information received from that, to derive confidence. Certainly in this task, when we asked people about the nature of others' recognition memory, they used statements about thoughts and feelings to produce categorisations of experience that are in keeping with actual first-person experiences.

In terms of single- and dual-process models of memory, these data are interpreted as being inconsistent with single-process models that posit that subjective experience and confidence are equivalent (e.g., Dunn, 2004). Regarding the current data, it is suggested that if recognition memory processes were best understood in terms of confidence (and thus by a single underlying process) then the effect of confidence on assignment of items to response categories should have been more comparable across categories of subjective experience. Contrary to this prediction, when confidence manipulations influenced categorisation of items to subjective experience categories in Experiment 2, the influence of confidence differed across item types. However, our data do not provide a definitive answer regarding the relationship between confidence and

subjective experience and we acknowledge that advocates of single-process viewpoints may draw different interpretations from our data.

One limitation of the current experimental approach is that the task we asked participants to perform was essentially one of pattern classification: we selected the stimuli; we showed participants our definitions of R, K, F, and G; and participants assigned the stimuli to those categories; i.e., there is some circularity in our design. One might suggest that the instruction to categorise statements according to our definitions was communicated to participants in such a manner that recovering the results we found was certain. This concern motivated our manipulation of confidence in Experiments 2 and 3, where it is not clear that our design would have led to shifts in classification according to confidence, since we were careful not to bias participants' judgments by including any mention of confidence in definitions of R, K, F, and G (see Table 1). As intrinsic confidence of items (assessed in Experiment 1) interfered with assignment of items to their original category of subjective experience (Experiment 2), we trust that our categories of subjective experience reflect natural conceptualisations of the relationship between phenomenological experiences and confidence in memory that people understand because of their first-person metacognitive experiences. However, this idea needs further support, and we plan to pursue investigation into natural conceptualisations of self and others' states of memory awareness in the future.

5.2 The split of Know into Know and Familiar

Our results demonstrate that Know and Familiar can be reliably differentiated by people both in terms of the confidence associated with the justification (Experiment 1) and, more importantly, in terms of some inherent differences in the justification statements, which lead to different patterns of categorisation (Experiments 2 and 3); although again we must note the

possibility of circularity here. That Know and Familiar are associated with different levels of confidence extends previous findings from recognition experiments to understandings of how people interpret others' memory experiences (Brewer & Sampaio, 2006; Dewhurst et al., 2009; and Kihlstrom et al., 1996). However, the confidence associated with Know judgments influenced how they were categorised. In Experiment 2, a large proportion of Know items were assigned to Familiar if they were Medium or Low confidence items and in Experiment 3, medium-confidence Know items paired with any level of confidence were equally likely to be assigned to Familiar as to Know. As discussed in Section 3.2.4, it is suggested that this categorisation stems from the high-confidence-without-recollection conceived as the basis of Know subjective experiences.

If Know and Guess experiences are related to confidence in one's evaluation of what is in memory then it follows that these responses would be influenced to a greater extent by confidence level when confidence is manipulated experimentally. The current patterns of data support the idea that Know and Familiar subjective experiences are dissociable and validates the use of the four separate categories of subjective experience in episodic memory paradigms. In sum, whereas the subjective experience of Familiarity appears to reflect the process of familiarity, the subjective state of Knowing appears to reflect high-confidence-without-recollection.

The findings of the current experiments concur with Gardiner et al.'s conclusion: "The contents of any particular mental experience are idiosyncratic. But the states of awareness are lawful" (Gardiner et al., 1998, p. 10). The novel finding of these experiments is that individuals' understanding of others' memory experiences is also lawful.

Acknowledgements

This research was undertaken as part of the first author's PhD research at the University of Leeds funded by a University Research Scholarship. The work derives from discussions at the ESRC-CNRS collaborative workshops programme awarded to Chris Moulin, (Recollection, Remembering and the Complex Nature of the Self; RES-170-25-0008). The authors wish to thank Jane Berry, Stephen Lindsay, and Mike Masson for their advice during the writing of this manuscript and Colleen Parks and two anonymous reviewers for their constructive comments on previous versions..

References

Barber, S. J., Rajaram, S., & Marsh, E. J. (2008). Fact learning: How information accuracy, delay, and repeated testing change retention and retrieval experience. *Memory, 16,* 934-946.

Brewer, W. F., & Sampaio, C. (2006). Processes leading to confidence and accuracy in sentence recognition: A metamemory approach. *Memory, 14,* 540-552.

Conway, M. A., Gardiner, J. M., Perfect, T. J., Anderson, S. J., & Cohen, G. M. (1997). Changes in memory awareness during learning: The acquisition of knowledge by psychology undergraduates. *Journal of Experimental Psychology: General, 126,* 393-413.

Dewhurst, S. A., Conway, M. A., & Brandt, K. R. (2009). Tracking the R-to-K shift: Changes in memory awareness across repeated tests. *Applied Cognitive Psychology, 23,* 849-858.

Diana, R. A., Reder, L. M., Arndt, A., & Park, H. (2006). Models of recognition: A review of arguments in favor of a dual-process account. *Psychonomic Bulletin & Review, 13,* 1-21.

Donaldson, W. (1996). The role of decision processes in remembering and knowing. *Memory & Cognition, 24,* 523-533.

Donaldson, W., MacKenzie, T. M., & Underhill, C. F. (1996). A comparison of recollective memory and source monitoring. *Psychonomic Bulletin & Review, 3,* 486-490.

Dunn, J. C. (2004). Remember-know: A matter of confidence. *Psychological Review, 111,* 524-542.

Dunn, J. C. (2008). The dimensionality of the remember-know task: A state-trace analysis. *Psychological Review, 115,* 426-446.

Gardiner, J. M. (2001). Episodic memory and autonoetic consciousness: a first-person approach. *Philosophical Transactions of the Royal Society of London Series B - Biological Sciences, 356,* 1351-1361.

Gardiner, J. M. (2008). Remembering and knowing. In J. H. Byrne (Ed.), *Learning and memory: A comprehensive reference* (pp. 285-305). Oxford: Elsevier Ltd.

Gardiner, J. M., & Java, R. I. (1990). Recollective experience in word and nonword recognition. *Memory & Cognition, 18,* 23-30.

Gardiner, J. M., Ramponi, C., & Richardson-Klavehn, A. (1998). Experiences of remembering, knowing, and guessing. *Consciousness and Cognition, 7*, 1-26.

Gardiner, J. M., Ramponi, C., & Richardson-Klavehn, A. (2002). Recognition memory and decision processes: A meta-analysis of remember, know, and guess responses. *Memory, 10*, 83-98.

Gardiner, J. M., & Richardson-Klavehn, A. (2000). Remembering and knowing. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 229-244). New York: Oxford University Press.

Gardiner, J. M., Richardson-Klavehn, A., & Ramponi, C. (1997). On reporting recollective experiences and "direct access to memory systems". *Psychological Science, 8,* 391-394.

Geraci, L., McCabe, D. P., & Guillory, J. J. (2009). On interpreting the relationship between remember-know judgments and confidence: The role of instructions. *Consciousness and Cognition, 18*, 701-709.

Herbert, D. M. B., & Burt, J. S. (2001). Memory awareness and schematization: Learning in the university context. *Applied Cognitive Psychology, 15*, 617-637.

Herbert, D. M. B., & Burt, J. S. (2003). The effects of different review opportunities on schematisation of knowledge. *Learning and Instruction, 13,* 73-92.

Herbert, D. M. B., & Burt, J. S. (2004). What do students remember? Episodic memory and the development of schematization. *Applied Cognitive Psychology, 18,* 77-88.

Kelley, C. M., & Jacoby, L. L. (1998). Subjective reports and process dissociation: Fluency, knowing, and feeling. *Acta Psychologica, 98*, 127-140.

Kihlstrom, J. F., Kim, M., & Dabady, M. (1996, November). *Remembering, knowing, and feeling in episodic recognition.* Paper presented at the Psychonomic Society Meeting, Chicago.

Macmillan, N. A., Rotello, C. M., & Verde, M. (2005). On the importance of models in interpreting remember-know experiments: Comments on Gardiner et al.'s (2002) meta-analysis. *Memory, 13*, 607-621.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review, 87*, 252-271.

McCabe, D. P., & Geraci, L. D. (2009). The influence of instructions and terminology on the accuracy of remember-know judgments. *Consciousness and Cognition, 18*, 401-413.

McCabe, D. P., Geraci, L., Boman, J. K., Sensenig, A. E., & Rhodes, M.G. (2011). On the validity of remember–know judgments: Evidence from think aloud protocols. *Consciousness and Cognition, 20*, 1625-1633.

O'Connor, A. R., Guhl, E. N., Cox, J. C., & Dobbins, I. G. (2011). Some memories are odder than others: Judgments of episodic oddity violate known decision rules. *Journal of Memory and Language, 64*, 299-315.

Parks, C. M., & Yonelinas, A. P. (2007). Moving beyond pure signal-detection models: Comment on Wixted (2007). *Psychological Review, 114*, 188-201.

Rajaram, S. (1993). Remembering and knowing: Two means of access to the personal past. *Memory & Cognition, 21*, 89-102.

Rajaram, S., Hamilton, M., & Bolton, A. (2002). Distinguishing states of awareness from confidence during retrieval: Evidence from amnesia. *Cognitive Affective & Behavioral Neuroscience, 2,* 227-235.

Rotello, C. M., Macmillan, N. A., & Reeder, J. A. (2004). Sum-difference theory of remembering and knowing: A two-dimensional signal-detection model. *Psychological Review, 111,* 588-616.

Rotello, C. M., Macmillan, N. A., Reeder, J. A., & Wong, M. (2005). The remember response: Subject to bias, graded, and not a process-pure indicator of recollection. *Psychonomic Bulletin & Review, 12,* 865-873.

Rotello, C. M., & Zeng, M. (2008). Analysis of RT distributions in the remember-know paradigm. *Psychonomic Bulletin & Review, 15,* 825-832.

Slotnick, S. D. (2010). "Remember" source memory ROCs indicate recollection is a continuous process. *Memory, 18,* 27-39.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology, 26,* 1-12.

Tunney, R. J., & Fernie, G. (2007). Repetition priming affects guessing not familiarity. *Behavioral and Brain Functions, 3,* 40.

Williams, H.L., Conway, M.A., & Moulin, C.J.A. (2013a). Pre-Exposure and source-discriminability: The use of recollection when items are familiar. Manuscript in preparation.

Williams, H.L., Conway, M.A., & Moulin, C.J.A. (2013b). The relationship between source, confidence, and subjective experience when two judgments are made post-recognition. Manuscript in preparation.

Williams, H.L., Moulin, C.J.A., & Conway, M.A. (2009). The use of recollection in associative recognition memory. *Proceedings of the Psychonomic Society, Vol. 14.*

Wixted, J. T., & Mickes, L. (2010). A continuous dual-process model of Remember/Know judgments. *Psychological Review, 117*, 1025-1054.

Wixted, J. T., & Stretch, V. (2004). In defense of the signal detection interpretation of remember/know judgments. *Psychonomic Bulletin & Review, 11*, 616-641.

Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 1341-1354.

Yonelinas, A. P. (2001). Consciousness, control, and confidence: The 3 Cs of recognition memory. *Journal of Experimental Psychology: General, 130*, 361-379.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46*, 441-517.

Yonelinas, A. P., Dobbins, I., Szymanski, M. D., Dhaliwal, H. S., & King, L. (1996). Signal-detection, threshold, and dual-process models of recognition memory: ROCs and conscious recollection. *Consciousness and Cognition, 5*, 418-441.