University of Calgary

PRISM Repository **https://prism.ucalgary.ca**

The Vault Open Theses and Dissertations

2018-01-05

Examining Memory States for General Knowledge: A Non-Trivial Pursuit

Pereverseff, Rosemary

Pereverseff, R. S. (2018). Examining Memory States for General Knowledge: A Non-Trivial Pursuit (Master's thesis, University of Calgary, Calgary, Canada). Retrieved from https://prism.ucalgary.ca.

http://hdl.handle.net/1880/106321

Downloaded from PRISM Repository, University of Calgary

UNIVERSITY OF CALGARY

Examining Memory States for General Knowledge: A Non-Trivial Pursuit

by

Rosemary Pereverseff

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN PSYCHOLOGY

CALGARY, ALBERTA

JANUARY, 2018

© Rosemary Pereverseff 2018

Abstract

According to the classic distinction between semantic and episodic memory, people answer general-knowledge questions by accessing their semantic memory. However, an appeal of trivia games is the variety of memory and metamemory experiences they arouse—which sometimes include recollection of episodic details. I report an in-depth exploration of subjective memory states for general knowledge. In two experiments, participants classified their answers for general-knowledge questions as *learning memory* or *related memory* forms of recollection, or as *feels familiar*, *just know*, or *guess* forms of non-recollection. Surprisingly, participants often reported recollection for their answers. Accuracy of answers assigned to the learning and related memory states, and to the just know state, were equally high. In contrast, accuracy was much lower for the feels familiar state (and was lowest for guesses). The accuracy difference between just know and feels familiar suggests these states are distinct, even though researchers often use them interchangeably. Reports of learning memories increased on an immediate retest. Recollection source judgments on the retest revealed that, in addition to recollecting the answer feedback provided on the initial test, pre-experimental recollections increased. Episodic memory is commonly experienced in semantic memory tasks and is diagnostic of accuracy.

Keywords: recall; metamemory; general knowledge; accuracy; recollection

Acknowledgements

I would like to thank my supervisor, Dr. Glen Bodner, for his guidance and expertise throughout my thesis. I would also like to thank the rest of my committee members, Drs. Andrea Protzner, Joshua Bourdage, and Jo Towers for offering their time and support throughout the development and review of my research.

Additionally, I would like to extend a thank you to my fellow graduate students, Alison Heard, David Sidhu, Devon Currie, as well as my lab mates, Shaela Jalava and Taylor Blanchette, who contributed their time in various ways to assist with my research.

Finally, I am grateful to my family and friends for their unconditional support.

Table of Contents

Abstract ii
Acknowledgements iii
Table of Contents iv

List of Tables vi
List of Figures and Illustrations vii

Chapter 1: Introduction 1
- **Using General-Knowledge Questions to Study Metamemory** 2
- **The Remember/Know Distinction in Metamemory** 5

Chapter 2: Experiment 1 10
- **Method** 12
  - **Participants** 12
  - **Materials** 12
  - **Memory States** 12
  - **Procedure** 14
- **Results and Discussion** 17
  - **Interrater Reliability for Classifying Memory State Descriptions** 17
  - **Memory States for Correct Answers** 19
  - **Memory States for Incorrect Answers** 23
  - **Expected Memory States for Answers** 25
  - **Accuracy of Memory States** 26

Chapter 3: Experiment 2 30
- **Method** 32
  - **Participants** 32
  - **Materials and Memory States** 32
  - **Procedure** 32
- **Results and Discussion** 35
  - **Replication of Experiment 1** 35
  - **Correct Answers in Recall** 35
  - **Incorrect Answers in Recall** 36
  - **Expected Memory States** 38
  - **Accuracy of Recall** 38
  - **Accuracy on the Multiple-Choice Questions** 39
  - **Recollection Source Judgments for Learning Memories** 41
  - **Confidence Measures** 45
  - **Hypercorrection** 47

Chapter 4: General Discussion 51
- **Drawing Distinctions Between Memory States for General Knowledge** 53
- **Erasing a Distinction Between Episodic and Semantic Memory?** 59
- **The Source of Learning Memories During a Retest** 61

Limitations and Future Directions......62
Implications and Conclusions......63

References......65

List of Tables

Table 1. Memory States for Answers to General-Knowledge Questions. 13

List of Figures and Illustrations

Figure 1. Experiment 1 procedure. 14

Figure 2. Mean number of correctly recalled (left) and incorrectly recalled (right) items assigned to each memory state in Experiment 1 (top), Experiment 2 initial test (middle) and Experiment 2 retest (bottom). Error bars indicate the standard error of the mean. Asterisks on the retest figures denote a significant difference ($p < .01$) in recall relative to the initial test for a given memory state. 18

Figure 3. Left panels show how memory states for shown answers were classified as a function of the memory state assigned to initial correct (top), incorrect (middle), or unanswered (bottom) recall in Experiment 1. Right panels show mean proportion of initial correct (top), incorrect (middle), and unanswered (bottom) recall for each memory state that matched the memory state for shown answers (with the number of participants contributing to each mean shown within each bar). Error bars indicate the standard error of the mean. 20

Figure 6. Experiment 2 procedure for initial test and retest phases. Dashed arrows indicate condition-specific parts of the procedure. 33

Figure 7. Mean number of correctly recognized (left) and incorrectly recognized (right) multiple-choice items assigned to each memory state on the initial test in Experiment 2. Error bars indicate the standard error of the mean. Asterisks denote a significant difference ($p < .01$) in recognition relative to recall on the initial test for a given memory state. 40

Figure 8. Accuracy of multiple-choice answers as a function of memory state (top), and accuracy of multiple-choice answers as a function of expected memory state for unanswered items (bottom) in Experiment 2. Error bars indicate the standard error of the mean. 42

Figure 9. Bar heights indicate the recollection source(s) selected for correct answers assigned to the learning memory state on the retest. Bar partitions indicate the initial test outcome for each set of items. Error bars indicate standard error of the mean for the entire bar height. (Pre = Pre-experiment memory; Pre+FB = Pre-experiment memory first when both were chosen; AllPre1st = Total times pre-experiment memory was chosen first; FB = Feedback memory; FB+Pre = Feedback memory first when both were chosen; AllFB1st = Total times feedback memory was chosen first). 43

Figure 10. Mean confidence rating as a function of memory state across correct and incorrect answers in Experiment 2 initial test (top) and retest (bottom). Error bars indicate the standard error of the mean. 47

Figure 11. Proportion correct on the Experiment 2 retest as a function of confidence in incorrect answers on the initial test. Error bars indicate the standard error of the mean. 49

Figure 12. Bar heights indicate the number of items assigned to each memory state on the retest for correct answers (left) and incorrect answers (right). Bar partitions indicate the initial test outcome for each memory state on the retest. Error bars indicate standard error of the mean for the entire bar height. 50

Chapter 1: Introduction

Testing one's memory for trivia—interesting but inessential facts—is a very popular pursuit. More than 100 million copies of the board game *Trivial Pursuit*® have been sold in 26 countries and in at least 17 languages since its creation in 1979 (Chilton, 2014). More than 2900 bars and restaurants in North America provide their patrons the opportunity to test their knowledge of trivia by playing *NTN Buzztime* (formerly *National Trivia Network*; Buzztime, 2016). And the television quiz show *Jeopardy!*, in which contestants are given a category and a trivia fact, and must answer in the form of a question, has endured in popularity for decades. Why do people take such interest in trying to remember obscure facts about topics ranging from pop culture to sports to geography? One motivation may be the rich and robust set of metamemorial experiences that people encounter when searching through their memories.

Metamemory refers to one's thoughts and feelings about one's memory (Dunlosky, Mueller, & Thiede, 2016; Flavell & Wellman, 1977). Trivia games arouse many types of metamemorial experiences. For example, take a moment try to generate an answer to the following general-knowledge question: *What is the longest river in South America?* In doing so, you may find yourself experiencing a *tip-of-the-tongue* (TOT) state in which you are sure you know the answer but cannot retrieve it. Alternatively, you might confidently answer "the Nile", even though that answer is incorrect. Or you might answer "the Amazon", but feel as though you are merely guessing, even though this answer is correct. Moreover, when the answer is revealed you might experience a range of reactions, from denial ("that's not a river in South America!") to euphoria (if a correct answer wins your team a trivia tournament). Indeed, it is not uncommon for game players to argue for an incorrect answer when presented the correct answer. In the present era, people often feel compelled to google the answers to general-knowledge questions that they cannot reconstruct from memory (e.g., Ferguson, McLean, & Risko, 2015).

Memory researchers have long used general-knowledge questions to study TOT states (e.g., Brennan, Baguley, Bright, & Bruce, 1990; Schwartz, 2001; 2008) in which an answer feels imminent but cannot be recalled. Surprisingly, however, researchers have not examined in detail people's memory states for the answers that they *are* able to generate. What subjective feelings do these answers arouse? How accurate are those feelings? How do those feelings shift when presented with the answers? How do those feelings shift when the same questions are answered again later on? The goal of the present study was to provide the first systematic study of memory and metamemory for answers to general-knowledge questions—and thus to provide answers to these and other questions.

Using General-Knowledge Questions to Study Metamemory

General-knowledge questions have been used to study a range of topics including intelligence (e.g., Furnham & Chamorro-Premuzic, 2006), memory in clinical and neuropsychological populations (e.g., Wojcik, Moulin, & Souchay, 2013, and Chua & Ahmed, 2016, respectively), the effects of aging on memory (e.g., Backman & Nilsson, 1996), and the effects of mood on memory and on confidence (e.g., Allwood & Bjorhag, 2008). Unlike traditional memory tasks that involve a study phase and a test phase, studies of memory for general knowledge do not typically involve a study phase. Rather, a person's lifetime serves as the study phase, so to speak.

Additionally, general-knowledge questions have often been used in the study of metamemory. Much of this research has examined the relationship between metamemory and later memory performance. *Prospective metamemory measures* are those collected during a study experience or after a person's memory fails (Nelson & Narens, 1990). These measures are often used to determine how well people are able to predict their later recognition. For example, participants are sometimes asked after a study trial to make a *judgment of learning* (JOL), that is, to rate how likely they think they will be to later recognize that item on a given memory test. JOLs are often predictive of later recognition, at least when there is a delay between study and test (Dunlosky & Nelson, 1992). However, because JOLs are collected during the study phase, they are not normally collected when the stimuli are general-knowledge questions.

Another prospective metamemory measure is a *feeling-of-knowing* (FOK) judgment. Unlike JOL ratings, FOK ratings are collected at test whenever participants fail to recall an answer (Souchay, Moulin, Clarys, Taconnat, & Isingrini, 2007). General-knowledge questions have long been used to study FOK states. For example, Hart (1965) used FOKs to examine whether participants can access information in memory about an answer despite being unable to retrieve it. Afterwards, the questions were presented again, this time on a multiple-choice test. Participants were more likely to select the correct answer on the latter test if they had indicated a FOK on the initial test. Further, Morson, Moulin, and Souchay (2015) used a test-study-retest design with general-knowledge questions to study FOKs with younger and older adults. A semantic task (test with no study phase) and an episodic task (study phase followed by a test) each consisted of a recall and multiple-choice portions, however they were separated by a study phase. Again, recognition accuracy was found to be higher after an FOK was indicated. Overall, FOKs have found to be indicative of memory for an item when used with general-knowledge questions.

General-knowledge questions have also often been used to study TOTs (Brown, 1991; Brown & McNeill, 1966). In a TOT state, one can often report structural details about the answer despite being unable to retrieve it (e.g., the letter it starts with, how many syllables it contains).

Using general-knowledge questions, Schwartz (2001) found that TOT states were predictive of retrieval time, where having a TOT state led to participants spending more time attempting to retrieve the target relative to not having a TOT state. In asking general-knowledge questions about celebrities (e.g., What actor played *Bionic Man*?), Brennan et al. (1990) showed that presenting the celebrities' initials helped participants resolve TOT states.

Not surprisingly, FOK ratings and reports of TOT states are correlated (Yaniv & Meyer, 1987), and both predict later recognition for answers (e.g., Brown, 1991; Metcalfe, Schwartz, & Joaquim, 1993; Nelson, 1984). However, the type of information accessed when making each judgment remains a matter of debate. For FOK judgments, one issue is whether they are based on partial knowledge (e.g., Hicks & Marsh, 2002; Koriat, 1993; 1995) and/or on familiarity for the cues presented in the question (e.g., Metcalfe et al., 1993; Reder & Ritter, 1992). In contrast, TOT states have been attributed to feelings of fluency (ease of accessibility as gauged by processing time) generated for non-retrieved answers (e.g., Cleary & Claxton, 2015). However, in common, FOKs and TOTs both focus on metamemory for the answers that participants cannot generate, which tells us less about people's metamemory for the answers they can generate. The present study measured metamemory both for questions participants can answer and for those they cannot answer.

Retrospective metamemory measures are collected after a participant makes a recognition or recall response at test (Nelson & Narens, 1990). One very common measure is a confidence rating collected during a test (e.g., Bornstein & Zickafoose, 1999; Koriat, 2008). Here, the confidence-accuracy relationship has been of particular interest. Of particular relevance to my thesis is the study of how initial confidence ratings on a test influence performance on a later memory test. Butterfield and Metcalfe (2001) reported that high-confidence errors for general-knowledge questions are corrected more often than low-confidence errors on later tests of memory, termed a *hypercorrection effect*. This effect is robust for answers to general-knowledge questions (Butler, Karpicke, & Roediger, 2008; Butterfield & Metcalfe, 2001, 2006; Fazio & Marsh, 2009; Sitzman, Rhodes, & Tauber, 2014; Sitzman, Rhodes, Tauber, & Liceralde, 2015). It has also been found that retention improves for low-confidence correct answers when feedback is provided (Butler et al., 2008). Taken together, these findings suggest that surprising feedback improves retention, perhaps by attracting increased attention (Butterfield & Metcalfe, 2006; Fazio & Marsh, 2009). However, as previously mentioned, other research suggests that confidence ratings might be based more on fluency of the questions/cues rather than partial memory for details of the answer (Koriat, 2008; McCabe & Soderstrom, 2011). As described below, my thesis tested whether metamemory judgments that direct participants to attend to the qualitative aspects of their memory also yielded a hypercorrection effect on a memory retest.

The Remember/Know Distinction in Metamemory

A different type of retrospective metamemory measure that is often collected at test involves asking participants to distinguish between experiences of *remembering* (recognition accompanied by recollection of episodic details) and experiences of *knowing* (recognition without recollection of episodic details). This measure has most often been collected during recognition tests, and has been used with materials as diverse as picture-word pairs (Dudukovic & Knowlton, 2006), university course materials (Conway, Gardiner, Perfect, Anderson, & Gillian, 1997; Herbert & Burt, 2004), and rare words and their definitions (Dewhurst, Conway, & Brandt, 2009).

Although remember/know memory states have most often been collected during recognition tests (i.e., as retrospective metamemory judgments), McCabe and Soderstrom (2011) had participants make prospective *judgments of remembering and knowing* (JORKs) at study. McCabe and Soderstrom (2011) suggested that JOLs rely on an evaluation of the overall fluency with which an item is processed at study, whereas JORKs might be based on an evaluation of the details of the encoding experience for an item. They predicted, and found, that the correlation with memory on a cued recall test was higher for JORKs than for JOLs.

To date, only one other study has collected prospective remember/know judgments. Hicks and Marsh (2002) had participants study cue-target word pairs (e.g., garage-camera) before taking a cued recall test (e.g., garage-???). When participants could not generate a target, they indicated whether they would have a remember, know, or guess experience when presented the answer on an immediate multiple-choice question. Hicks and Marsh found that prospective remember/know/guess judgments during the cued recall test predicted subsequent recognition accuracy.

Finally, while remember/know judgments have primarily been used with recognition tests, Mickes, Seale-Carlisle, and Wixted (2013) collected them during a free recall test. Surprisingly, they found that many recalled items were classified as “know” experiences, even though participants had presumably used recollection to recall them. As detailed below, building on these few studies, my thesis is the first to examine the recollection-based and non-recollection-based memory states that arise for answers to general-knowledge questions.

The distinction between *remembering* and *knowing* was proposed by Tulving (1985), who suggested that remember/know judgments made by participants during a recognition test might reveal whether items were retrieved from episodic memory, leading to *remember* responses, or from semantic memory, leading to *know* responses. Additionally, the act of retrieving information from each of these memory systems was theorized by Tulving to pertain to different types of consciousness. According to this theory, awareness of memory obtained from episodic memory, and the experience of remembering, is correlated with *autonoetic consciousness*, whereas that obtained from semantic memory, and the experience of knowing, is correlated with *noetic consciousness*. Although Tulving first conceptualized the remember-know task under the episodic-autonoetic/semantic-noetic distinctions, these alternative concepts have been used interchangeably to describe the dual process account of metamemory.

Much debate has ensued regarding whether the *remember* and *know* memory states reflect separate underlying memory processes (e.g., Yonelinas, 2002), such as those discussed above, or different levels of memory strength of a single underlying process (e.g., Donaldson, 1996; Dunn, 2004; 2008; Wixted & Mickes, 2010). Remember judgments have been shown to have higher levels of confidence associated with it than know judgments (Donaldson, 1996; Dunn, 2004; 2008), supporting a single process account. However, they have also been shown to have more source-memory details than know responses, even when controlling for strength/accuracy (Wixted & Mickes, 2010). This evidence provides support for Tulving's (1985) dual-process account, and suggests that these two memory states pick out qualitatively and subjectively different retrieval experiences.

Debate has also arisen around a possible distinction between two non-recollection memory states: *knowing* and *familiarity*. *Familiarity* (recognition accompanied by a nagging feeling without recollection of episodic details) has often been used interchangeably with *knowing* in the metamemory literature. However, a number of researchers have provided evidence that these two states can be differentiated by participants (Barber, Rajaram, & Marsh, 2008; Conway et al., 1997; Dewhurst et al., 2009), alluding to an important underlying difference. In more closely examining these differences, Williams and Moulin (2014) found that

that non-expert participants divided others' justifications for using know responses into two types that reliably mapped onto experts' definitions for familiarity and knowing. Thus, based on this line of research, these two memory states reflect qualitatively different subjective experiences.

Because Tulving's (1985) notion of noetic consciousness pertains to information that comes to mind involuntarily, presumably from semantic memory where information about the world is categorically stored, memory research has most commonly assumed that general-knowledge questions, when not preceded by a study phase, tap into a person's semantic memory. For example, Blaxton (1989) states that "both the general knowledge and the fragment completion tests were classified as semantic memory tasks because both could be performed without reference to the original study episode…" (p. 659). As another example, Souchay et al. (2007) used general-information questions to tap semantic memory. If general-knowledge questions do not invoke episodic memory and autonoetic consciousness then there would not be much utility in using the remember/know paradigm with such questions. Perhaps as a result of this assumption, the remember/know paradigm has not been used to date to examine memory for general-knowledge questions.

And yet, in the context of a trivia game, both remember and know experiences commonly occur. Consider the trivia question posed earlier. A player in a trivia game might recall learning about the Amazon river from a geography teacher in middle school. This would constitute a remember/episodic/autonoetic experience, because the memory is connected to the person's personal past. However, another player might provide the same answer but be unable to bring to mind any details of where, when, or how this fact was learned. This would constitute a know/semantic/noetic experience. According to Tulving (1972), these two players might be accessing the answer information from different memory systems (i.e., episodic vs. semantic memory, respectively). However, one need not make a separate memory system claim to examine this distinction (cf. Aue, Criss, & Prince, 2015). From a metamemory perspective, the key is that these two types of experience feel different to people, and as a result, people may systematically distinguish between them. In the present study, I carried out two experiments with the aim to determine the frequency and accuracy with which remember and know judgments are experienced for general-knowledge questions.

Chapter 2: Experiment 1

In Experiment 1, participants attempted to answer a set of general-knowledge questions. They then retrospectively classified their memory state as one of two recollection state—*learning memory* or *related memory*—as one of two non-recollection states—*feels familiar* or *just know*—or as a non-memory-based state—*guess*. Recollection was separated into two different types on an exploratory basis into learning memories (recollection of the context in which an answer was learned) and related memories (recollection of information related to the answers but not of learning the answer itself). When piloting people's memory for answers to general-knowledge questions, they often distinguished these two types of recollection. Because remember/know judgments are traditionally based on memory for a study phase, the definition for remembering refers to that specific learning event. But with general-knowledge questions, related episodic details might constitute a distinct experience.

Additionally, based on literature discussed above that indicates that familiarity and knowing are two distinct types of non-recollection experiences (Barber et al., 2008; Conway et al., 1997; Dewhurst et al., 2009; Williams & Moulin, 2014), I provided both options to participants. Finally, I included the guess option as a non-memory-based state because previous research has suggested that the guess option is based on strategies other than accessing directly relevant memories (Gardiner, Ramponi, and Richardson-Klavehn, 1998).

When participants could not answer a question, participants indicated the memory state they expected to experience upon viewing the correct answer (Hicks & Marsh, 2002; McCabe & Soderstrom, 2011). This constituted a prospective metamemory measure. In all cases, answer feedback was then provided. Participants then indicated their memory state for the shown answer. Experiment 1 thus aimed to explore the distribution of memory states for correct, incorrect, and unanswered questions, as well as the accuracy of answers among the memory states, and the stability of each memory state in the face of feedback.

Method

Participants

University of Calgary undergraduates (n = 77; mean age = 20.2; female = 65) participated for course credit. Participants indicated fluency in English. Four additional participants were excluded for not following instructions.

Materials

The materials were 76 general-knowledge questions (e.g., *What is the longest river in South America?* Answer: *Amazon*) drawn from Tauber, Dunlosky, Rawson, Rhodes, and Sitzman (2013). Their mean probability of correct recall was .42 (*SD* = .25; range = .08-.84) to ensure there were enough correct answers for analysis, and ideally enough incorrect answers and/or expected memory states for analysis as well.

Memory States

As outlined in Table 1, participants were offered 4 memory state responses and 1 non-memory-state response designed to capture the types of metamemory experiences that might arise when answering general-knowledge questions: *learning memory*, *related memory*, *feels familiar*, *just know,* and *guess*. The traditional recollection memory state was divided into two separate categories. *Learning memory* referred to what is classically thought of as recollection of episodic details about the learning event. The *related memory* category was novel, and referred to the experience of recollecting details of an event related to the answer, while not recollecting the original learning event itself. The episodic details retrieved during a related memory experience are not about the origin of one's memory for the answer itself.

Table 1. Memory States for Answers to General-Knowledge Questions.

Memory State	Definition and Example
Learning memory	When the answer came to mind, you also remembered a specific prior experience of learning this fact (e.g., how, when, and/or where you learned it). It does not have to be a memory of the first time you learned it. *For example, you might remember learning that the cheetah is the fastest land animal from Ms. Jones in Grade 3 when you were doing a project about Africa, or from a top 10 list that you read online last week.*
Related memory	When the answer came to mind, you also remembered a specific related personal memory, but the memory was not about learning this fact. *For example, you might remember playing with a toy cheetah with your cousin, or watching an episode of a TV show about cheetahs on the savannah, but this related memory was not specifically about learning the fact that the cheetah is the fastest land animal.*
Feels familiar	When the answer came to mind, a specific personal memory did not come to mind. Instead, the answer just seemed very familiar to you. *For example, the answer cheetah "feels familiar" to you.*
Just know	When the answer came to mind, a specific personal memory did not come to mind. Instead, the answer just seems to be part of your general knowledge. *For example, you "just know" that the world's fastest land animal is cheetah.*
Guess	You do not experience a specific personal memory, a feeling of familiarity, or a feeling of knowing for your answer. It is just an educated guess. *For example, you are just guessing that the fastest land animal might be the cheetah.*

The two non-recollective memory states were feels familiar and just know. Neither of these memory states have episodic details associated with them. The *feels familiar* state occurs when one's answer feels familiar (e.g., the knowledge that Mother's Day is in May might feel familiar), whereas the *just know* state occurs when the answer comes to mind but with no particular feeling state associated with it (e.g., the knowledge that Christmas Day is December 25th is just known). Finally, a *guess* option was included to allow participants to report an answer as an educated guess for which they had no memory or feeling state.

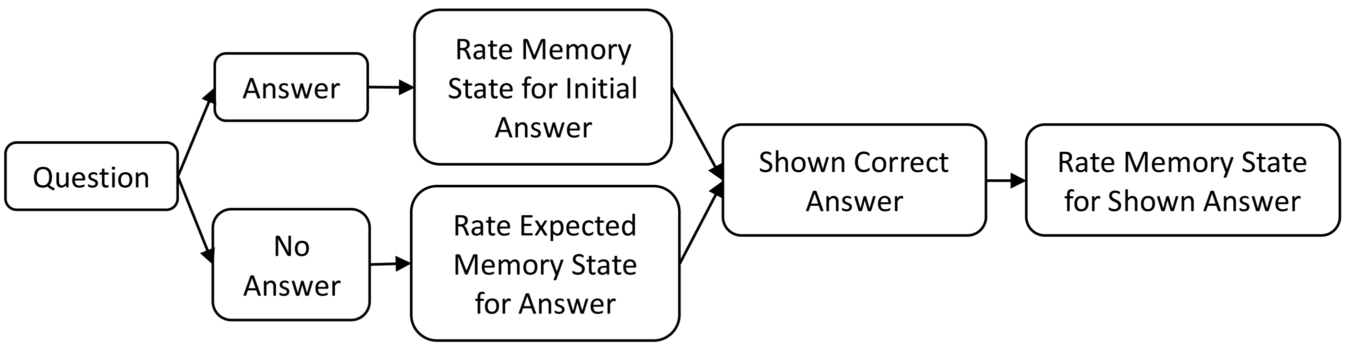

Figure 1. Experiment 1 procedure.

Procedure

The study design is presented in Figure 1. Participants were tested online through Qualtrics survey software. At the start of the experiment, participants were asked to focus on the experiment and to not complete other activities. Participants were informed that they would be shown a series of general-knowledge questions one at a time (e.g., *What is the longest river in South America?*), that they would spend up to 10 seconds trying to generate an answer for each question, and that they would then type their answer into a response box. They were further told that if they could not think of an answer, they should type in "@". To prevent accidental entries participants had to enter either an answer or “@”. Participants were then told that they would be asked to classify their memory for their initial answer into 1 of 5 memory states. If they could not come up with an answer, they were told that they would instead rate their expected memory state for the answer for the same 5 memory states. An expected memory state was defined as the memory state they thought they would experience when they were shown the correct answer. They were then told that they would be shown the answer, and would classify their memory for it from 1 of the 5 states (regardless of whether their initial answer was correct). Participants were then given the definitions and examples of the 5 memory states as shown in Table 1. They were told that they would then be tested for their understanding of each memory state.

To ensure that participants could distinguish the 5 memory states, participants were then shown an example of each memory state that they needed to correctly classify. For example, for the question, “Manga are a type of comic that originates in what country?”, the example memory state was, “When the answer Japan comes to my mind, I remember learning that it is the answer while watching Jeopardy with my roommate”, to which the correct answer, “learning memory”, would have to be chosen. This procedure was repeated for all 5 memory states in a random order. When participants had correctly identified each memory state, they moved on to the test phase. They were reminded that they should try to answer as many questions as they could, making educated guesses where possible. The instructions were presented again for them to review. The 76 general-knowledge questions were then presented in an order randomized for that participant. When selecting their memory state or expected memory state, the definitions and examples appeared at the bottom of the page for reference, and participants clicked the corresponding button on the screen.

Afterwards, participants were asked whether there were any questions where they thought that their memory state did not fit into any of the provided categories (this was almost never the case), and how often they played trivia games (*very rarely, rarely, occasionally, frequently,* or *very frequently*; the modal response was *rarely*). Finally, a memory state description question asked participants the following question for each memory state:

> Please provide an example of an answer you rated as “[memory state]” and describe why you chose “[memory state]” for this answer. If you cannot remember an example, please describe in general why you sometimes chose “[memory state]” as your memory state.

The experiment typically took about 40 minutes to complete.

Results and Discussion

Results were significant at $p < .05$ except where Bonferroni corrected. Partial eta-squared (η_p^2) is reported as a measure of effect size. For repeated-measures ANOVAs, Greenhouse-Geisser corrections were used to adjust for violations of sphericity. Misspellings were scored as correct if they could be unambiguously interpreted (e.g., “rasin” for “raisin”), but answers that spelled another name or word were scored as incorrect (e.g., “marina” for “Mariana”).

Interrater Reliability for Classifying Memory State Descriptions

Each memory state was used at least once by 96% of participants showing that most participants distinguished between the learning and related memory states, and also between familiarity and just knowing states. Two independent raters were provided with each participant’s memory state descriptions for the two recollection states arranged in a random order (i.e., either the learning or related memory first). The raters coded one response as the learning memory and the other as the related memory. Their accuracy rates were 96% and 95%, and there was very good agreement between the two raters, $\kappa = .87$ (Altman, 1991). Thus, the raters were able to distinguish participants’ use of the learning versus related memory states. The same two raters also performed a task coding each participant’s “feels familiar” and “just know” descriptions. Their accuracy rates were 93% and 96%, and there was again very good agreement between the two raters, $\kappa = .78$. Thus, the raters were also able to distinguish participants’ use of the feels familiar versus just know memory states. Taken together, these accuracies demonstrate that raters were able to pick out the distinctions that the participants made between pairs of states. However, they do not guarantee that participants made valid and appropriate distinctions between these pairs of memory states that were based on the provided definitions. I elaborate on this issue in the General Discussion.

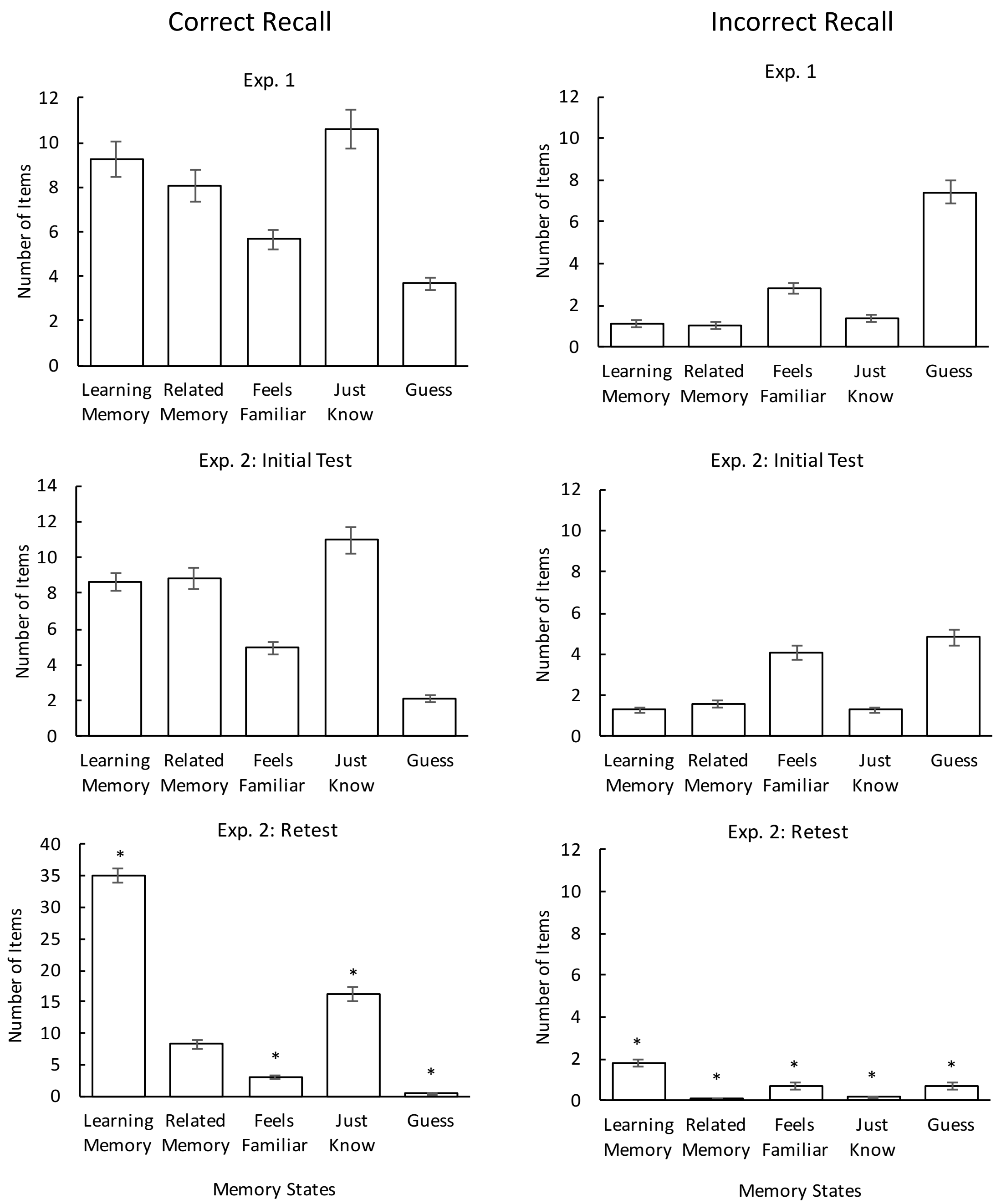

Figure 2. Mean number of correctly recalled (left) and incorrectly recalled (right) items assigned to each memory state in Experiment 1 (top), Experiment 2 initial test (middle) and Experiment 2 retest (bottom). Error bars indicate the standard error of the mean. Asterisks on the retest figures denote a significant difference ($p < .01$) in recall relative to the initial test for a given memory state.

Memory States for Correct Answers

The mean proportion of correct initial answers was .49 (37.4 items), slightly higher than the .42 mean expected based on the Tauber et al. (2013) norms. Figure 2 (top left) shows the mean number of correct items assigned to each memory state. A repeated-measures ANOVA revealed a significant difference in the frequency of reporting of the 5 memory states, $F(2.767, 210.297) = 17.17$, $MSE = 50.79$, $\eta_p^2 = .18$, $p < .001$. Post-hoc comparisons (Bonferroni adjusted $p = 0.007$) revealed that the following memory states were reported similarly often for correct answers: learning memory vs. related memory (9.27 vs. 8.10), $F(1,76) = 1.22$, $p = .27$, learning memory vs. just know (9.27 vs. 10.64), $F(1,76) = 1.06$, $p = .31$, and related memory vs. just know (8.10 vs. 10.64), $F(1,76) = 4.49$, $p = .04$. Each of these states was reported more often than the feels familiar state: learning memory vs. feels familiar (9.27 vs. 5.66), $F(1,76) = 16.11$, $MSE = 62.29$, $\eta_p^2 = .18$, $p < .001$, related memory vs. feels familiar (8.10 vs. 5.66), $F(1,76) = 10.19$, $MSE = 45.07$, $\eta_p^2 = .12$, $p < .001$, and just know vs. feels familiar (10.64 vs. 5.66). $F(1,76) = 24.50$, $MSE = 77.76$, $\eta_p^2 = .24$, $p < .001$. Finally, the feels familiar state was reported more often than the guess state (5.66 vs. 3.68), $F(1,76) = 14.86$, $MSE = 20.46$, $\eta_p^2 = .16$, $p < .001$. Thus, a substantial number of correct answers to general-knowledge questions generated recollection experiences; in fact, each recollection state was reported as often as the just know state. This outcome is surprising given Tulving's (1972) notion that general-knowledge questions tap semantic memory. Additionally, the finding that correct answers were almost twice as likely to be classified as just know than as feels familiar suggests that participants were able to make a distinction between these two memory states that researchers have often used interchangeably.

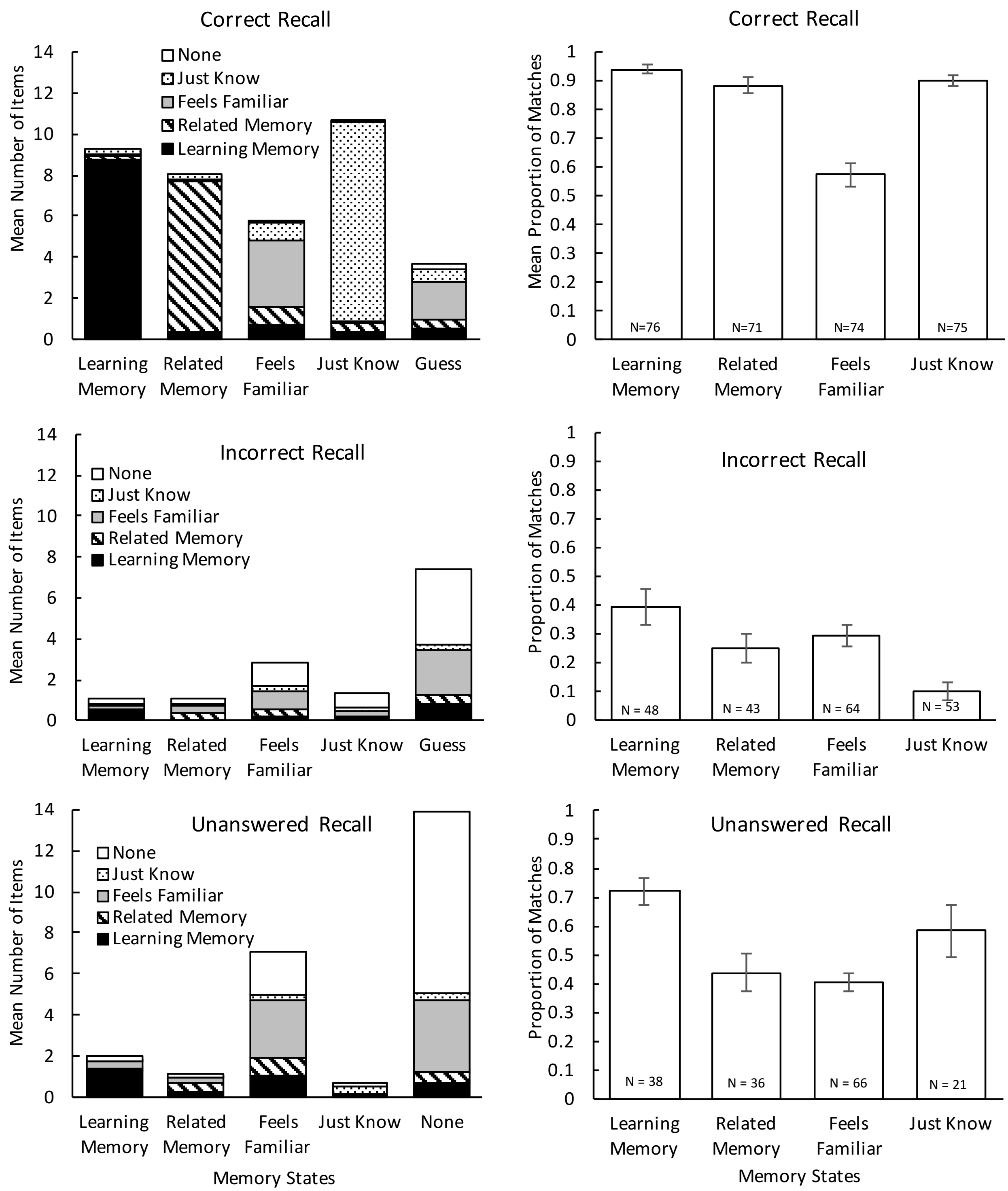

Figure 3. Left panels show how memory states for shown answers were classified as a function of the memory state assigned to initial correct (top), incorrect (middle), or unanswered (bottom)

recall in Experiment 1. Right panels show mean proportion of initial correct (top), incorrect (middle), and unanswered (bottom) recall for each memory state that matched the memory state for shown answers (with the number of participants contributing to each mean shown within each bar). Error bars indicate the standard error of the mean.

The bar partitions in Figure 3 (top left) illustrate how, after providing a correct answer, participants then classified their memory state for the shown answer. Participants' memory states typically did not change when shown the answer (which constituted confirmatory feedback for correct answers). Correct guesses, which had to shift states because a guess option was not provided for the shown answers, most often shifted to the feels familiar state. Thus, despite learning that their initial answer was correct, most participants who indicated that their memory state for their initial answer lacked recollection (i.e., feels familiar, just know, or guess) continued to not report experiencing recollection for the shown answer. In other words, participants appear to have "taken their best shot" at recollecting the answer at the outset. There was little indication of a hindsight bias (Fischhoff, 1975), in which participants fall prey to a bias of claiming to recollect answers once they are shown to them. Potentially, then, correct initial answers that were not recollected—either initially or after viewing the answer—represent cases of retrieval from semantic memory.

I next calculated the proportion of memory states for shown answers that matched the initial memory state reported for correct initial answers by using the means and bar partition means shown in Figure 3 (top left). For example, if a participant indicated related memory for 10 of their correct answers, and then indicated related memory for 8 of these 10 answers when shown the answers, then their proportion match for the related memory state would be .80.

Guesses were excluded from analysis because participants could not choose guess for the shown answer.

Figure 3 (top right) presents the mean proportion of matches for each memory state. A repeated-measures ANOVA compared these 4 proportions of matching states (excluding 11 participants who did not assign 1 or more correct answers to each state). The ANOVA showed a significant difference in the proportion of matches, $F(2.19, 142.44) = 42.16$, $MSE = 0.06$, $p < .001$. Post-hoc follow-up contrasts (Bonferroni adjusted $p = .007$) showed that the proportion of matches was high and similar for learning memory vs. related memory (.94 vs. .88), $F(1,65) = 5.54$, $MSE = 0.05$, $p = .02$, for related memory vs. just know (.88 vs. .91), $F(1,65) = 0.83$, $MSE = 0.07$, $p = .37$, and for learning memory vs. just know (.94 vs. .91), $F(1,65) = 1.52$, $MSE = 0.04$, $p = .22$. Thus, these 3 states were equally stable in the presence of correct answer feedback. In contrast, the proportion of matches for the learning memory, related memory, and just know states were each higher than for the feels familiar state (.57), $F(1,65) = 74.67$, $MSE = 0.12$, $p < .001$, $F(1,65) = 48.22$, $MSE = 0.13$, $p < .001$, and $F(1,65) = 54.70$, $MSE = 0.14$, $p < .001$, respectively. Thus, the feels familiar state was particularly labile in the face of answer feedback. Feels familiar responses for an initial answer were similarly likely to shift to a learning memory, a related memory, or a just know state for the answer (.24 vs. .21 vs. .25; $N = 74$), $F < 1$. The nagging feeling that accompanies the feels familiar state was frequently an intermediary memory state which, when resolved in the presence of confirmatory answer feedback, transformed with similar probability into one of the other memory states. This transformation of the memory state experienced for a correct answer occurred much less often for answers associated with the just know state, providing additional evidence (beyond the raters' ability to distinguish participants' reports) that the feels familiar and just know states represent different experiences.

Memory States for Incorrect Answers

The mean proportion of incorrect answers was .18 (13.8 items). Figure 2 (top right) shows the mean number of incorrect answers assigned to each memory state. A repeated-measures ANOVA revealed a significant difference in the frequency of reporting the 5 memory state options for incorrect answers, $F(1.792, 136.156) = 77.64$, $MSE = 16.19$, $\eta_p^2 = 0.51$, $p < .001$. Post-hoc comparisons (Bonferroni adjusted $p = .007$) showed that the following memory states were reported similarly infrequently for incorrect answers: learning memory vs. related memory (1.17 vs. 1.08), $F < 1$, learning memory vs. just know (1.17 vs. 1.38), $F(1,76) = 1.59$, $p = .21$, and related memory vs. just know (1.08 vs. 1.38), $F(1,76) = 1.42$, $p = .24$. The feels familiar state was reported more often than each of these states: feels familiar vs. learning memory (2.82 vs. 1.12), $F(1,76) = 32.95$, $MSE = 6.77$, $\eta_p^2 = 0.30$, $p = .001$, feels familiar vs. related memory (2.82 vs. 1.08), $F(1,76) = 35.39$, $MSE = 6.59$, $\eta_p^2 = 0.32$, $p < .001$, and feels familiar vs. just know (2.82 vs. 1.38), $F(1,76) = 18.62$, $MSE = 8.59$, $\eta_p^2 = 0.20$, $p < .001$. The guess state was assigned to incorrect answers more often than any other memory state, as shown when comparing guess vs. feels familiar (the next most frequent state) (7.43 vs. 2.82), $F(1,76) = 52.92$, $MSE = 30.93$, $\eta_p^2 = 0.41$, $p < .001$. Thus, incorrect answers were rarely experienced as either type of recollection state or as just know, suggesting that participants used these memory state categories carefully and/or did not often experience false recollection or false knowing.

Figure 3 (middle left) shows that, after offering initial answers that were incorrect, participants typically experienced the shown answers as either feels familiar or as eliciting no memory state. When incorrect answers were initially offered as guesses, the memory state reported for the shown answer (which had to shift away from guess) was usually no memory ($M = 4.0$ items), or feels familiar ($M = 2.3$ items). After offering incorrect initial answers, shown answers were rarely classified as recollected, again suggesting that participants did not fall prey to a hindsight bias when shown the answers.

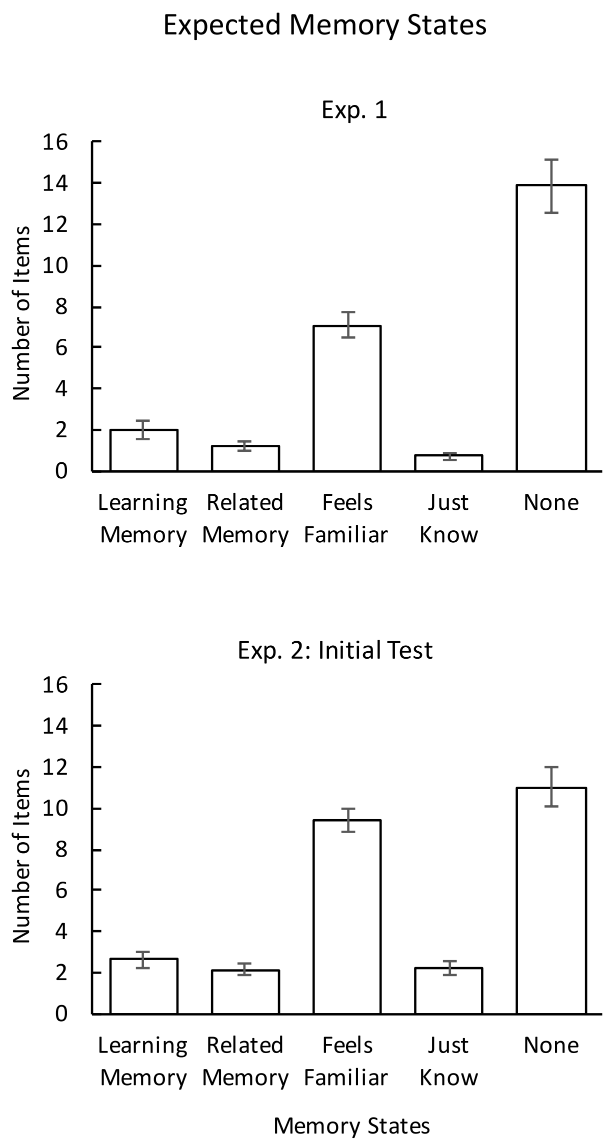

Figure 4. Mean number of unanswered items in Experiment 1 (top) and Experiment 2 initial test (bottom) assigned to each expected memory state.

The mean proportion of matching memory states, for each memory state, are shown in

Figure 3 (middle right). For example, if a participant indicated feels familiar for 10 of their incorrect initial answers, then indicated feels familiar for 6 of these 10 answers when shown the answers, then their proportion match for the feels familiar state would be .60. These means were not analyzed because only 14 participants attributed at least 1 incorrect answer to each initial memory state.

Expected Memory States for Answers

The mean proportion of unanswered questions was .33 (24.8 items). Figure 4 (top) shows the memory states participants expected for the answers to these questions. A repeated-measures ANOVA revealed significant differences among the number of responses for each expected memory state, $F(1.674, 127.192) = 66.33$, $MSE = 86.94$, $\eta_p^2 = 0.47$, $p < .001$. Post hoc comparisons (Bonferroni adjusted $p = .007$) revealed that the following memory states were reported similarly often for unanswered questions: learning memory vs. related memory (1.99 vs. 1.17), $F(1,76) = 3.72$, $p = .06$, and related memory vs. just know (1.99 vs. 0.70), $F(1,76) = 2.85$, $p = .096$. However, the learning memory state was expected more often than the just know state (1.99 vs. 0.70), $F(1,76) = 8.00$, $MSE = 15.92$, $\eta_p^2 = 0.10$, $p = .006$. The feels familiar state was expected more often than each of these states: feels familiar vs. learning memory (7.10 vs. 1.99), $F(1,76) = 40.77$, $MSE = 49.45$, $\eta_p^2 = .35$, $p < .001$, feels familiar vs. related memory (7.10 vs. 1.17), $F(1,76) = 70.82$, $MSE = 35.77$, $\eta_p^2 = .50$, $p < .001$, and feels familiar vs. just know (7.10 vs. .70), $F(1,76) = 84.57$, $MSE = 37.32$, $\eta_p^2 = 0.53$, $p < .001$. This difference in the expectation of the feels familiar and just know memory states once again highlights an important difference between the two states. Finally, participants expected to experience no memory state more often than any other state, as shown when comparing this option to feels familiar (the next most frequent state) (13.87 vs. 7.10), $F(1,76) = 19.34$, $MSE = 182.23$, $\eta_p^2 = 0.20$, $p < .001$. Thus, participants most often expected to experience no memory state for the answer, followed by a feels familiar state. Participants did not generally expect to experience either recollection state, thus they rarely experienced a classic tip-of-the-tongue state (e.g., Brown, 1991) in which they felt expected to experience recollection for the shown answer.

The bar partitions in Figure 3 (bottom left) show the memory states participants reported for the shown answers for these items. In general, participants who did not generate an initial answer typically experienced either no memory state (M = 8.8 items) or the feels familiar state (M = 2.8 items) for the shown answers, in line with their expectations. This concordance could be genuine or could reflect a bias to report the memory state that they expected to have.

The mean proportion of matching memory states, for each memory state, are shown in Figure 3 (bottom right). For example, if a participant indicated just know state for 10 of their expected memory states, then indicated just know for 5 of these 10 for the shown answer, their proportion of matches would have been .50. Only 9 participants classified at least 1 unanswered question into each memory state, thus precluding analysis.

Accuracy of Memory States

For each memory state, an accuracy score was calculated by dividing the number of correct answers assigned to a given state by the number of correct and incorrect answers assigned to a given state (i.e., correct/correct+incorrect). The mean accuracy across the 76 questions was .73, which was significantly higher than the .59 accuracy for these questions in the Tauber et al. (2013) norms, $t(76) = 10.13$, $SE = .013$, $p < .001$. Differences in the samples, testing procedure (e.g., requirement to provide memory states and expected memory states), data scoring, or even the test context (i.e., the subset of items selected from the norms) might account for this difference.

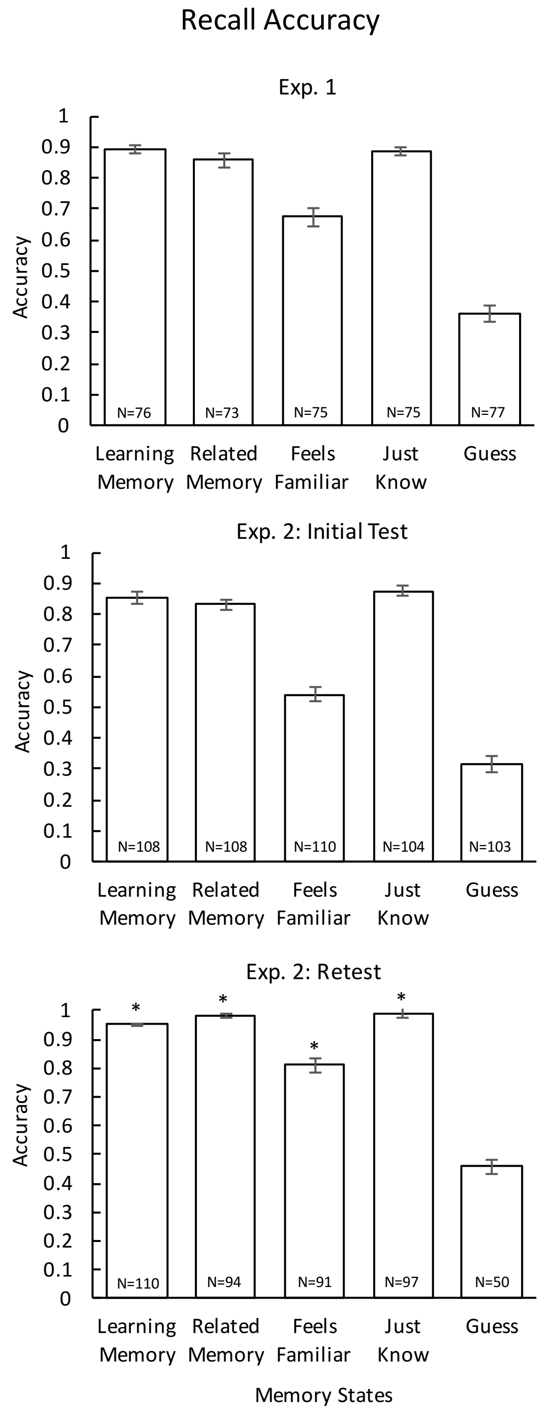

Figure 5. Accuracy of recall for each memory state in Experiment 1 (top), Experiment 2 initial test (middle), and Experiment 2 retest (bottom). Error bars indicate the standard error of the mean.

The mean accuracy for each memory state is shown in Figure 5 (top). The pattern for accuracy was very similar to that for the distribution of memory states for correct answers (Figure 2, top left) given the low rate of incorrect answers. A repeated-measures ANOVA (excluding 9 participants who did not classify at least 1 correct or 1 incorrect answer to each memory state) indicated a difference in accuracy across the 5 states, $F(3.225, 216.055) = 116.24$, $MSE = .04$, $\eta_p^2 = .63$, $p < .001$. Post-hoc comparisons (Bonferroni adjusted $p = .008$) showed that accuracy was high and similar for the following pairs of memory states: learning memory vs. related memory (.89 vs. .86), $F(1,67) = 1.31$, $MSE = .04$, $p = .25$, learning memory vs. just know (.89 vs. .89), $F < 1$, and related memory vs. just know (.86 vs. .89), $F < 1$. The similarity in accuracy between the two recollection states is surprising; it could indicate that recollecting a specific learning memory for an answer does not improve accuracy over recollecting a related memory that is not about the answer itself. I consider this possibility, along with others, in the General Discussion. Additionally, experiencing recollection did not increase accuracy over and above just knowing, implying that recollection is not a more trustworthy memory state than just knowing in the predictive accuracy sense. Instead, it provides additional evidence that recollection and just knowing reflect subjectively different but similarly accurate memory states.

The just know state was more accurate than the feels familiar state (.89 vs. .68), $F(1,67) = 48.89$, $MSE = .06$, $\eta_p^2 = .42$, $p < .001$. This difference provides further support for the notion that there is an important distinction between these two states, and should not be used interchangeably as has often been the case in the metamemory literature. The related memory state was also more accurate than the feels familiar state (.86 vs. .68), $F(1,67) = 27.88$, $MSE = .08$, $\eta_p^2 = .29$, $p < .001$. However, the feels familiar state was more accurate than the guess state (.68 vs. .37), $F(1,67) = 87.45$, $MSE = .07$, $\eta_p^2 = .57$, $p < .001$. Thus, answers associated with the feels familiar state were more accurate than answers associated with the guess state, but were less accurate than answers associated with the other memory states.

Chapter 3: Experiment 2

Experiment 1 revealed that recollective forms of memory were commonly experienced when answering general-knowledge questions. In addition, the feels familiar and just know memory states were dissociable in terms of their use and accuracy. Moreover, memory states for correct answers were largely stable in the face of feedback. Experiment 2 was designed as a replication of Experiment 1 but with four key additions.

First, I examined whether participants whose initial answers were incorrect, or who could not provide an answer, could recognize the correct answer. To this end, a multiple-choice question was added after each recall attempt, rather than having participants rate their memory state for the shown correct answer. This protected against the possibility that participants would show hindsight bias in which once they were shown the answer they might be tempted to report a recollection memory state for it.

Second, I determined how the distribution and accuracy of memory states changed after corrective feedback, and whether a hypercorrection effect can be observed using qualitative memory state reports (Butterfield & Metcalfe, 2001). To this end, an immediate retest phase was added (following Butterfield & Metcalfe, 2001, 2006, and Morson et al., 2015). I expected that learning memories would increase on the retest relative to the initial test because the answer feedback would still be fresh in participants' memories.

Third, and building on the last point, the retest was also used to determine whether answer feedback might enhance participants' ability to recollect general knowledge from memory. To explore this novel question, on the retest, when participants reported a recollection state for an item, they were asked to indicate the source of their memory (i.e., feedback or pre-experiment). Past research using other types of materials has shown that viewing answer feedback can lead to a shift in memory states from knowing to remembering (Dudukovic & Knowlton, 2006; Dewhurst et al., 2009).

Fourth, I collected confidence ratings for one group of participants after they reported their memory state. This was done in part to further explore whether confidence was higher for the more accurate just know state than for the less accurate feels familiar state. If so, this would suggest that these two non-recollective states might differ quantatively rather than qualitatively. Additionally, following the hypercorrection literature (e.g., Butterfield & Metcalfe, 2001), collecting confidence ratings allowed me to see whether I could obtain a hypercorrection effect between confidence on the initial test and accuracy on the retest. Past research has shown that accuracy and confidence for memory for general knowledge are correlated (Brewer & Sampaio, 2012; Koriat, 2008). Based on this research, I expected that both recollection states, as well as the just know state, would have high confidence, based on their similar levels of accuracy.

Method

Participants

Additional University of Calgary undergraduates (*N* = 110; mean age = 21.1; female = 93) from the Experiment 1 pool participated for course credit. They were randomly assigned to either the confidence rating condition (*n* = 57; mean age = 21.4; female = 49) or the no-confidence rating condition (*n* = 53; mean age = 20.7; female = 44). Participants indicated fluency in English. Two additional participants were excluded for not following instructions.

Materials and Memory States

The materials and memory states were identical to Experiment 1. To create the multiple-choice test options, 3 plausible incorrect alternative answers were chosen for each of the 76 general-knowledge questions. Tauber et al. (2013) provided common commission errors for a subset of our 76 questions, so these were used as alternatives when available. The remaining alternatives were drawn using the Latent Semantic Analysis (LSA) website (Laham & Steinhart, 1998), as well as WordNet (Princeton University, 2010). Seven colleagues were asked to circle plausible alternatives for each question and at least 1 person circled an incorrect alternative for each question.

Procedure

The study procedure is outlined in Figure 6. Participants were tested in the lab rather than being tested online (cf. Experiment 1). All instructions for the experiment were presented on the computer screen and were verbally summarized by the experimenter.

The initial test procedure was identical to Experiment 1, with two exceptions. First, after indicating their memory state for an answer, participants in the confidence rating condition rated their confidence in their answer from 0 (lowest) to 100 (highest). Second, rather than simply being shown the correct answer after rating their memory state, the question was provided again along with four alternatives, in multiple-choice question format. After choosing an answer, participants rated their memory state for it, and were then shown the correct answer.

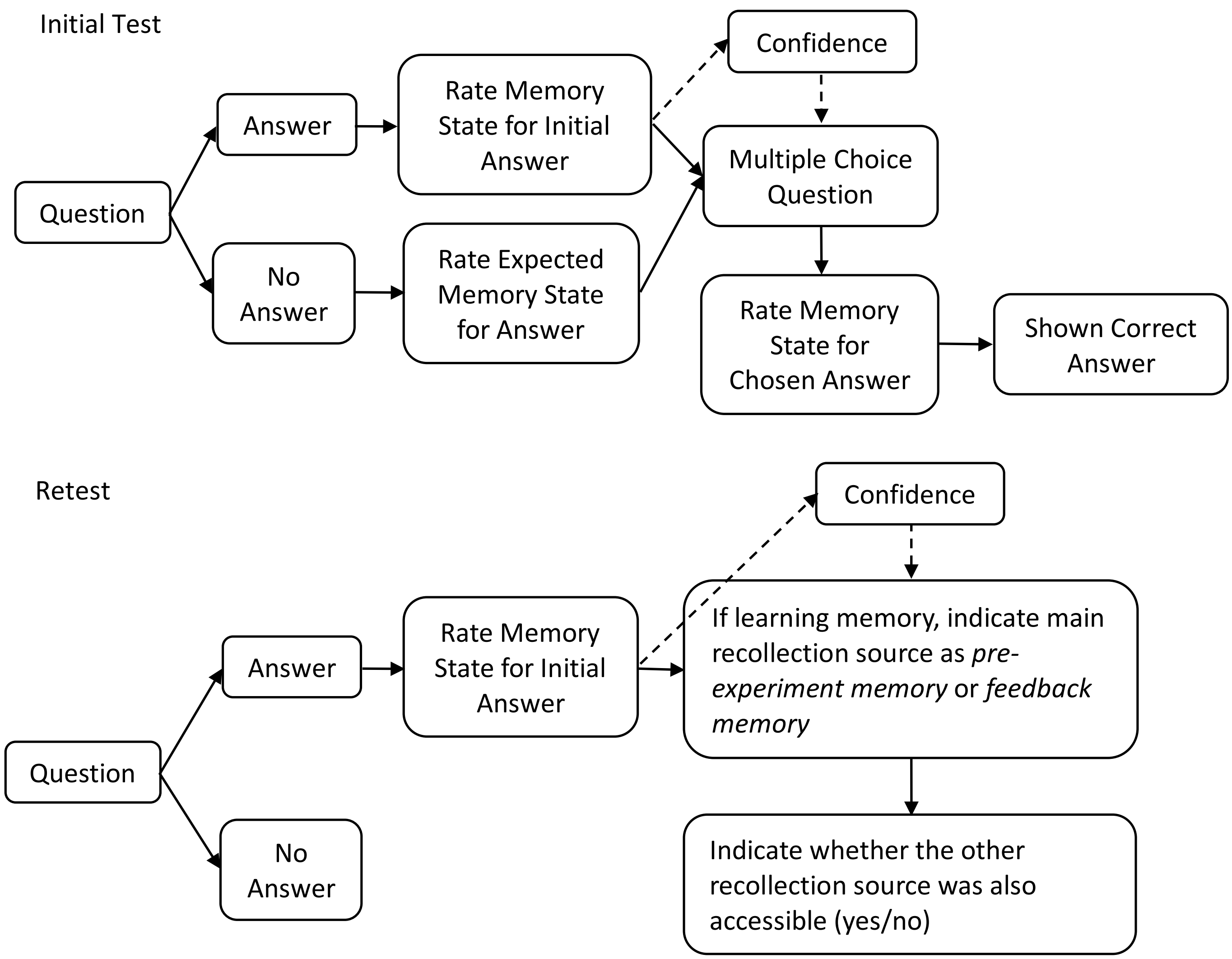

Figure 6. Experiment 2 procedure for initial test and retest phases. Dashed arrows indicate condition-specific parts of the procedure.

The retest phase began right after the initial test. The 76 general-knowledge questions were presented again in a re-randomized order, and participants attempted to answer each question and then indicated their memory state (and confidence, where applicable) for their answer, as in the initial test phase. However, if participants could not provide an answer, they moved on to the next question rather than rating their expected memory state.

Additionally, when the memory state for an answer on the retest was a learning memory, the following *recollection-source question* was then presented:

> "Was your reported answer based on memory of information learned prior to the study (*pre-experiment memory*) or on recollection of seeing the answer feedback on the initial test (*feedback memory*)?"

After responding, participants were then asked a follow-up question:

> "Can you also remember the answer from [*memory source they did not choose*]?"

This process was repeated for each of the 76 general-knowledge questions. Afterward, participants were asked whether there were any questions where they thought that their memory state did not fit into any of the provided categories (this was seldom indicated), and how often they played trivia games (the modal response was *rarely*). The session took approximately 1 hr.

Results and Discussion

Coding and analysis of the initial test results largely followed Experiment 1 except as noted. Importantly, the results were very similar for the confidence rating and no-confidence rating conditions, such that the reported analyses collapse across them unless otherwise stated.

Replication of Experiment 1

The initial test results in Experiment 2 were generally very similar to Experiment 1, including the mean number of correctly recalled answers (35.5 vs. 37.4 items), $F(1,185) = 1.18$, $MSE = 133.97$, $p = .28$, incorrect answers (13.1 vs. 13.8 items), $F < 1$, and unanswered questions (27.4 vs. 24.8 items), $F(1,185) = 1.93$, $MSE = 157.43$, $p = .17$. The distributions of memory states were also very similar (see Figures 2 and 4), as was the overall accuracy across the 76 questions here relative to Experiment 1 (.72 vs. .73). Thus, participants performed similarly to Experiment 1 despite the addition of confidence ratings (for participants in the confidence rating condition) and the inclusion of multiple-choice questions and answer ratings.

Correct Answers in Recall

The average number of correct answers assigned to each memory state on the initial test and on the retest are shown in Figure 2 (middle and bottom left). A 2 (test: initial vs. retest) x 5 (memory state: learning memory, related memory, feels familiar, just know, guess) repeated-measures ANOVA on correct recall was conducted. The main effect of test revealed, not surprisingly, that correct recall increased from the initial test to the retest ($M = .47$ vs. .83; 35.7 vs. 63.1 items), $F(1,109) = 1650.77$, $MSE = 5.09$, $\eta_p^2 = 0.94$, $p < .001$. The main effect of memory state was significant, $F(2.059, 224.409) = 147.46$, $MSE = 193.08$, $\eta_p^2 = 0.58$, $p < .001$, as was the interaction, $F(2.138, 233.039) = 278.05$, $MSE = 53.75$, $\eta_p^2 = .72$, $p < .001$.

The distribution of memory states on the initial test was much the same as in Experiment 1, despite the method changes in Experiment 2 (see Figure 2). On the retest, the major change was the large increase in correct answers attributed to the learning memory state due to the provision of answer feedback after the multiple-choice questions on the initial test. The interaction was followed up by comparing the number of items assigned to each memory state on the initial test vs. retest (Bonferroni adjusted $p = .01$). As just mentioned, the number of correct answers assigned to the learning memory state greatly increased from the initial test to the retest (8.62 vs. 35.06), $F(1,109) = 659.48$, $MSE = 58.33$, $\eta_p^2 = 0.86$, $p < .001$. The number of correct items assigned to related memory state did not change from initial test to retest (8.34 vs. 8.26), $F < 1$. This dissociation across initial test and retest for the learning memory and related memory states suggests that the two states may be distinct. Reports of just know states increased modestly from initial test to retest (10.95 vs. 16.22), $F(1,109) = 50.69$, $MSE = 30.16$, $\eta_p^2 = 0.32$, $p < .001$, whereas reports of the feels familiar state decreased from the initial test to the retest (4.95 vs. 3.12), $F(1,109) = 21.86$, $MSE = 8.40$, $\eta_p^2 = 0.17$, $p < .001$, as did reports of the guess states (2.14 vs. 0.45), $F(1,109) = 71.87$, $MSE = 2.17$, $\eta_p^2 = 0.40$, $p < .001$. Again, the dissociation across the two tests for the feels familiar and just know states may suggest that the two states represent distinct memory states. I return this issue in the General Discussion.

Incorrect Answers in Recall

The mean number of incorrect answers assigned to each memory state on the initial test and retest are shown in Figure 2 (middle and bottom right). In the ANOVA, a main effect of test revealed that the average number of incorrect items assigned to each memory state decreased from the initial test to the retest (2.62 vs. 0.69), $F(1,109) = 379.55$, $MSE = 2.69$, $\eta_p^2 = 0.78$, $p < .001$. A main effect of memory state was also present, $F(2.879, 313.857) = 35.26$, $MSE = 6.97$,

$\eta_p^2 = 0.24$, $p < .001$, as was a significant interaction, $F(2.856, 311.323) = 54.47$, $MSE = 4.78$, $\eta_p^2 = 0.33$, $p < .001$.

On the initial test, the distribution of incorrect recall among the memory states was similar to what was found in Experiment 1, however there was a major shift on the retest towards fewer incorrect answers overall. This shift was not surprising, given participants received answer feedback on the initial test. Post-hoc comparisons (Bonferroni adjusted $p = .01$) were conducted to follow up the significant interaction. The number of incorrect items assigned to learning memory increased marginally from initial test to retest (1.79 vs. 1.30), $F(1,109) = 5.26$, $p = .024$. This increase in learning memories for both correct and incorrect answers suggests that participants may have developed a response bias toward choosing the learning memory state on the retest. Alternatively, it may indicate the existence of false recollections of having viewed the incorrect answer or choosing it in the multiple-choice question. In either case, the increase in correct learning memories on the retest far exceeded the increase in incorrect learning memories, and, as shown below, the accuracy of the learning memory state increased on the retest. For related memories, the number of incorrect answers reported decreased from the initial test to the retest (1.61 vs. 0.11), $F(1,109) = 87.73$, $MSE = 1.41$, $\eta_p^2 = 0.45$, $p < .001$. Thus, as with correct recall, this dissociation in how the number of items assigned to learning and related memory states changed after viewing feedback highlights an important underlying difference between these two states. The numbers were also greater on the initial test than the retest for incorrect items assigned to feels familiar (4.06 vs. 0.70), $F(1,109) = 161.60$, $MSE = 3.85$, $\eta_p^2 = 0.60$, $p < .001$, just know (1.31 vs. 0.16), $F(1,109) = 77.24$, $MSE = 0.93$, $\eta_p^2 = .42$, $p < .001$, and guess (4.81 vs. 0.69), $F(1,109) = 122.54$, $MSE = 7.61$, $\eta_p^2 = 0.53$, $p < .001$. These decreases can likely be attributed to the overall decrease in the number of incorrect answers after viewing feedback.

Expected Memory States

Figure 4 (bottom) shows the number of unanswered questions assigned to each expected memory state. The pattern of expected memory states for the initial test in Experiment 2 was very similar to Experiment 1 (see Figure 4), and followed the pattern shown for incorrect answers. The ANOVA revealed a significant difference in the rate of expected memory states, $F(2.092, 227.998) = 61.47$, $MSE = 65.29$, $\eta_p^2 = 0.36$, $p < .001$. Post hoc comparisons (Bonferroni adjusted $p = .007$) indicated that the following memory states were equally low and similar: learning memory vs. related memory (2.62 vs. 2.15), $F(1,109) = 3.01$, $MSE = 8.16$, $p = .085$, learning memory vs. just know (2.62 vs. 2.21), $F < 1$, and related memory vs. just know (2.15 vs. 2.21), $F < 1$. Feels familiar was more likely to be expected than the following states: feels familiar vs. learning memory (9.41 vs. 2.62), $F(1,109) = 84.88$, $MSE = 59.76$, $\eta_p^2 = 0.44$, $p < .001$, feels familiar vs. related memory (9.41 vs. 2.15), $F(1,109) = 126.84$, $MSE = 45.76$, $\eta_p^2 = 0.54$, $p < .001$, and feels familiar vs. just know (9.41 vs. 2.21), $F(1,109) = 115.16$, $MSE = 49.52$, $\eta_p^2 = 0.51$, $p < .001$. This difference in reports of expected memory state between feels familiar and just know once again points to an important underlying difference underlying the two states. Finally, the number of items for which participants indicated the feels familiar state or no memory state were similarly high (9.41 vs. 11.04), $F(1,109) = 2.21$, $MSE = 131.80$, $p = 0.14$.

Accuracy of Recall

Figure 5 (middle and bottom) presents the mean accuracy for each memory state on the initial test and retest. The accuracy pattern on the initial test was similar to Experiment 1, and was also similar to the pattern on the retest. To examine how accuracy of the memory states shifted from initial test to retest, a 2 (test: initial vs. retest) x 4 (memory state: learning memory vs. related memory vs. feels familiar vs. just know) repeated-measures ANOVA was conducted (excluding 40 participants who did not classify at least 1 correct or incorrect answer to each memory state on both tests, and excluding the guess state due to its rarity of use on the retest). The ANOVA revealed a significant main effect of test, reflecting a sizeable increase in accuracy from the initial test to the retest (.79 vs. .95), $F(1,69) = 165.26$, $MSE = 0.02$, $\eta_p^2 = 0.71$, $p < .001$. There was also a main effect of memory state, $F(1.724, 118.98) = 65.63$, $MSE = 0.04$, $\eta_p^2 = 0.49$, $p < .001$, which was qualified by a significant interaction, $F(2.376, 163.946) = 20.13$, $MSE = 0.02$, $\eta_p^2 = .23$, $p < .001$. Pairwise tests (Bonferroni adjusted $p = .01$) showed that the accuracy increased across the initial and retest for each memory state: for the learning memory state (.85 vs. .95), $F(1,107) = 33.74$, $MSE = 0.02$, $\eta_p^2 = 0.24$, $p < .001$; for the related memory state (.84 vs. .98), $F(1,92) = 70.45$, $MSE = 0.01$, $\eta_p^2 = .43$, $p < .001$; for the feels familiar state (.55 vs. .81), $F(1,90) = 72.42$, $MSE = 0.04$, $\eta_p^2 = .45$, $p < .001$; for the just know state (.90 vs. .99), $F(1,92) = 81.46$, $MSE = .01$, $\eta_p^2 = .47$, $p < .001$. The increase for the guess state was not significant (.36 vs. .46), $F(1,48) = 2.86$, $MSE = 0.09$, $\eta_p^2 = 0.06$, $p = .097$. Thus, as in Experiment 1, accuracy was similarly high for learning memory, related memory, and just know states, was lower for the feels familiar state, and was lowest for guesses.

Accuracy on the Multiple-Choice Questions

In Experiment 1, participants were presented the answer after their recall attempt. In Experiment 2, a multiple-choice question was administered after each recall attempt during the initial test, allowing me to determine whether participants could pick out (recognize) the correct answer. Figure 7 shows the numbers of correct and incorrect multiple-choice answers assigned to each memory state. Figure 8 (top) shows the accuracy of multiple-choice answers as a function of the memory state reported for that answer, as well as a function of the expected memory state predicted when an answer was not provided (bottom). In both cases, the accuracy pattern for the memory states on the multiple-choice test (Figure 8) was very similar to the pattern for correct recall (Figure 2) and for recall accuracy (Figure 5).

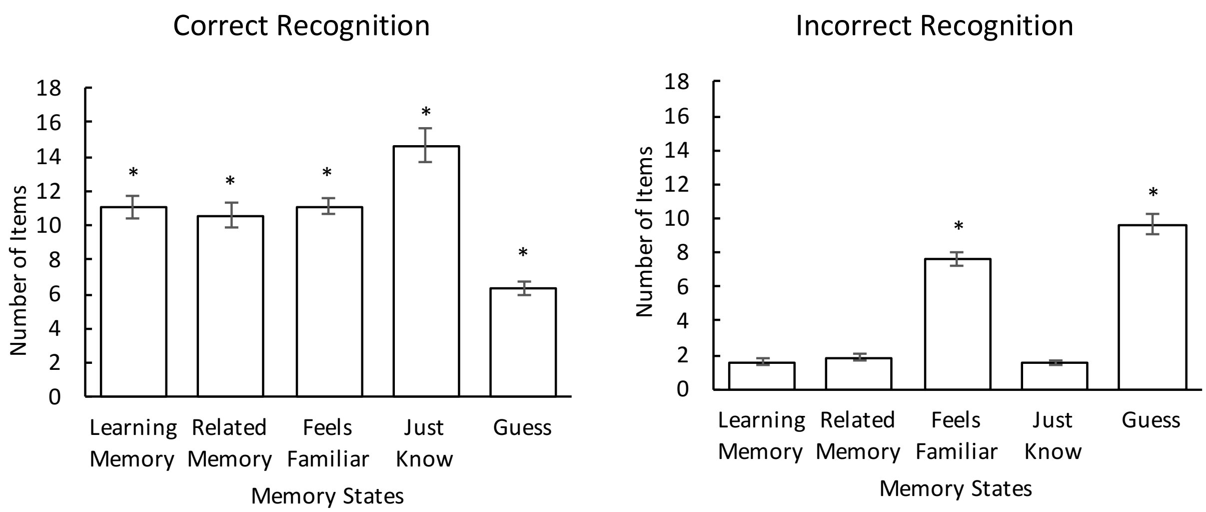

Figure 7. Mean number of correctly recognized (left) and incorrectly recognized (right) multiple-choice items assigned to each memory state on the initial test in Experiment 2. Error bars indicate the standard error of the mean. Asterisks denote a significant difference (p < .01) in recognition relative to recall on the initial test for a given memory state.

To further examine the accuracies on the multiple-choice questions as a function of the memory state reported for that answer, a repeated-measures ANOVA was conducted (excluding 5 participants who did not assign an answer to each memory state at least once). This analysis showed significant differences among these accuracies, $F(3.885, 404.047) = 288.49$, $MSE = 0.02$, $\eta_p^2 = 0.74$, $p < .001$. Pairwise comparisons (Bonferroni adjusted $p = .007$) revealed that multiple-choice accuracy was high and similar for the following memory states: learning memory vs. related memory (.88 vs. .85), $F(1,104) = 2.61$, $p = .109$, learning memory vs. just know (.88 vs. .89), F < 1, and related memory vs. just know (.85 vs. .89), $F(1,104) = 2.79$, $p = .098$. The accuracy for the following memory states were all greater than for the feels familiar state: learning memory vs. feels familiar (.88 vs. .61), $F(1,104) = 269.52$, $MSE = 0.03$, $\eta_p^2 = 0.72$, $p < .001$, related memory vs. feels familiar (.85 vs. .61), $F(1,104) = 194.09$, $MSE = 0.03$, $\eta_p^2 = 0.65$, $p < .001$, and just know vs. feels familiar (.89 vs. .61), $F(1,104) = 217.98$, $MSE = 0.04$, $\eta_p^2 = 0.68$, $p < .001$. Finally, the accuracy for guesses was lower than the accuracies for the other states, as evidenced by the comparison to the next lowest accuracy: guess vs. feels familiar (.40 vs. .61), $F(1,104) = 152.\ 32$, $MSE = 0.03$, $\eta_p^2 = 0.59$, $p < .001$. Thus, the patterns here once again match those found in Experiment 1 and on the recall portion of the initial test in Experiment 2.

The accuracies for the multiple-choice questions as a function of the expected memory states reported during the recall test was not analyzed due to the small number of observations, but the general trend was quite similar to that for correct answers (see Figure 8).

Recollection Source Judgments for Learning Memories

An important and novel aspect of my study was having participants indicate the recollected source(s) for the learning memories they experienced for their answers on the retest. The two learning memory sources defined on the retest were *pre-experiment*, which referred to learning memories that occurred prior to the start of the experiment, and *feedback*, which referred to recollecting learning the answer from the initial test phase. On the retest, participants who reported experiencing the learning memory state were first asked which of these sources their learning memory response was based on. They were then asked if they could also recollect the answer from the other source. Three sets of analyses were conducted on these recollection source judgments. The source selections for correct answers classified as learning memory on the retest are presented in Figure 9 (given the high accuracy on the retest, source selections for incorrect answers were not analyzed). Although it was quite common for participants to report having access to both pre-experiment and feedback recollections, most of the analyses reported below used the source they reported basing their learning memory response on (labeled "AllPre1st" and "AllFB1st" in Figure 9). This approach is in keeping with how source judgments are typically collected and analyzed.

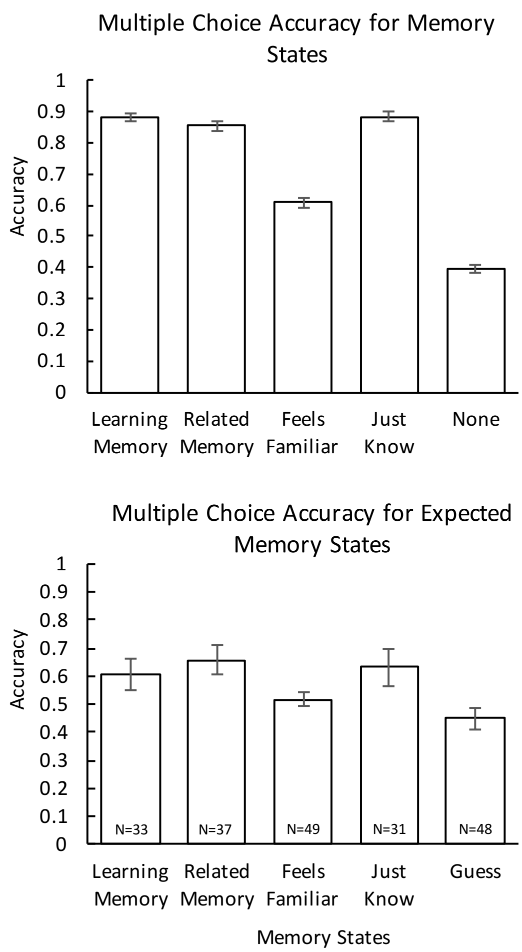

Figure 8. Accuracy of multiple-choice answers as a function of memory state (top), and accuracy of multiple-choice answers as a function of expected memory state for unanswered items (bottom) in Experiment 2. Error bars indicate the standard error of the mean.

The first analysis revealed that the majority of learning memories on the retest were recollections of the feedback on the initial test rather than pre-experiment recollections (25.0 vs. 10.1 items), $F(1,109) = 176.38$, $MSE = 69.06$, $\eta_p^2 = 0.62$, $p < .001$. This result was not unexpected, given that participants viewed the answer feedback just prior to the retest.

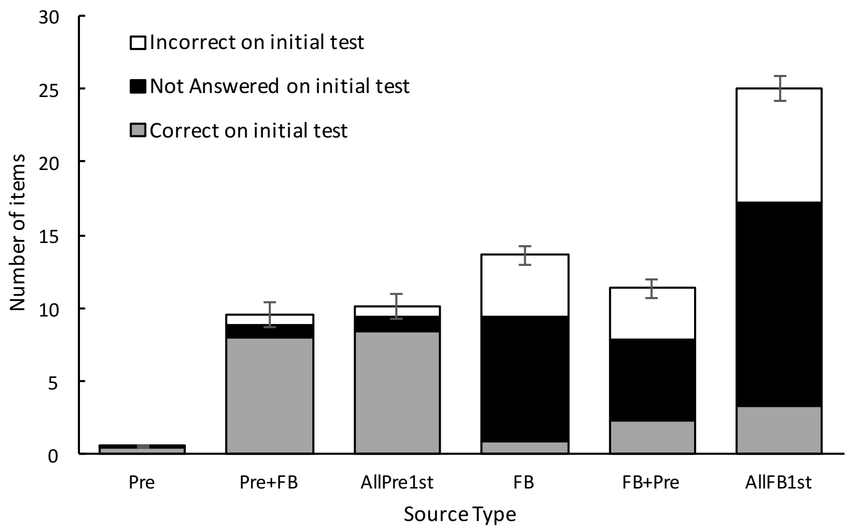

Figure 9. Bar heights indicate the recollection source(s) selected for correct answers assigned to the learning memory state on the retest. Bar partitions indicate the initial test outcome for each set of items. Error bars indicate standard error of the mean for the entire bar height. (Pre = Pre-experiment memory; Pre+FB = Pre-experiment memory first when both were chosen; AllPre1st = Total times pre-experiment memory was chosen first; FB = Feedback memory; FB+Pre = Feedback memory first when both were chosen; AllFB1st = Total times feedback memory was chosen first).

The second analysis examined whether the rate of reporting each recollection source depended on whether participants were correct, incorrect, or did not answer the question on the initial test (see bar partitions for the “AllPre1st” and “AllFB1st” bars in Figure 9). Because of issues of dependency, I computed the proportion of each source as a function of initial test performance (correct, incorrect, or unanswered), then conducted separate repeated-measures ANOVAs comparing each proportion across the two sources. Strikingly, a higher proportion of pre-experiment-based learning memories than feedback-based learning memories were correct on the initial test (.81 vs. .12), $F(1,107) = 714.61$, $MSE = 0.04$, $\eta_p^2 = 0.87$, $p < .001$. Conversely, a higher proportion of feedback-based learning memories than pre-experiment-based learning memories were incorrect on the initial test (.33 vs. .08), $F(1,107) = 126.06$, $MSE = 0.03$, $\eta_p^2 = 0.54$, $p < .001$, or unanswered on the initial test (.55 vs. .10), $F(1,107) = 327.22$, $MSE = 0.03$, $\eta_p^2 = 0.75$, $p < .001$. Thus, when participants correctly answered a question on the initial test, they tended to report pre-experiment learning memories on the retest, whereas when they were wrong on the initial test or could not answer the question then they tended to report feedback learning memories on the retest. Thus, feedback was particularly potent when the initial answer was incorrect, or when participants did not provide an answer.

The third analysis examined whether answer feedback on the initial test helped participants access pre-experiment recollections during the retest. If so, then relative to the number of learning memories on the initial test, there may be more pre-experiment learning memories on the retest. Conversely, the answer feedback might overshadow pre-experiment recollections during the retest, resulting in a decrease. Contrary to the latter possibility, there were more pre-experiment learning memories on the retest than learning memories on the initial test (10.09 vs. 8.62), $F(1,109) = 5.56$, $MSE = 21.45$, $\eta_p^2 = .05$, $p = .020$, though the magnitude of the difference was not striking (a gain of about 1 item). However, a more liberal analysis that included learning memories based on feedback for which a pre-experiment recollection also occurred (i.e., FB+Pre in Figure 9) revealed a large increase in the number of correct answers associated with pre-experiment recollection on the retest (21.46 vs. 8.62), $F(1,109) = 220.11$, $MSE = 41.23$, $\eta_p^2 = 0.67$, $p < .001$. Therefore, despite "taking their best shot" to recollect the answer as shown in Experiment 1, this novel result suggests that answer feedback enabled participants to recollect pre-experiment memories that they did not have access to during the initial test. Thus, not only is episodic memory involved in the recall of answers to general-knowledge questions, but viewing answer feedback may increase the number of pre-test episodic memories people experience on a subsequent test.

Confidence Measures

The mean confidence ratings (based on participants in the confidence-rating condition) for each memory state are shown in Figure 10. These means were based on ratings for both correct and incorrect answers because the pattern of means was very similar for both answer types and averaging across them resulted in less missing data. As with the accuracy analyses, confidence ratings for the guess state were excluded from analysis (fewer than half of the participants used the guess state on the retest). Thus, a 2 (test: initial vs. retest) x 4 (memory state: learning memory vs. related memory vs. feels familiar vs. just know) repeated-measures ANOVA was conducted (for which 18 participants were excluded due to not classifying at least 1 answer to each of the other 4 memory states). The ANOVA revealed a main effect of test, indicating an increase in confidence from initial test to retest (80.4 vs. 93.6), $F(1,38) = 110.14$, $MSE = 122.38$, $\eta_p^2 = 0.74$, $p < .001$. A main effect of memory state was also present, revealing a significant difference in confidence ratings across the 4 memory states, $F(1.387, 52.701) = 93.49$, $MSE = 253.32$, $\eta_p^2 = 0.71$, $p < .001$. As can be seen in Figure 10, the confidence pattern followed the accuracy pattern in Figure 5 (and the correct recall pattern in Figure 2).

Nonetheless, because this was the first time that confidence was measured, post-hoc comparisons (Bonferroni adjusted $p = .01$) were performed. Confidence ratings were high and nearly identical for answers assigned to the learning and related memory states (91.6 vs. 91.4), $F < 1$. However, the confidence ratings were even higher for the just know state than for the learning memory state (95.5 vs. 91.6), $F(1,38) = 21.30$, $MSE = 55.08$, $\eta_p^2 = 0.36$, $p < .001$, or the related memory state (95.5 vs. 91.4), $F(1,38) = 24.74$, $MSE = 53.24$, $\eta_p^2 = 0.39$, $p < .001$. Confidence ratings were much higher for each of these states than for the feels familiar state, including the related memory state (91.4 vs. 69.5), $F(1,38) = 88.86$, $MSE = 422.80$, $\eta_p^2 = 0.70$, $p < .001$. Importantly, confidence ratings were also much higher for the just know state than for the feels familiar state (95.5 vs. 69.5), $F(1,38) = 129.37$, $MSE = 409.32$, $\eta_p^2 = 0.77$, $p < .001$, even though the instructions avoided using definitions for these states that might indicate that confidence should be higher for the just know state. Thus, the just know state yields both higher accuracy and higher confidence than the feels familiar state, confirming that the two states are distinguishable. However, it does not rule out the possibility that the two states differ quantitatively rather than qualitatively. I return to this issue in the General Discussion.

Returning to the 2 x 4 ANOVA analysis, a significant interaction was also present, $F(1.706, 64.828) = 12.61$, $MSE = 102.47$, $\eta_p^2 = 0.25$, $p < .001$. This interaction was followed up by comparing confidence ratings on the initial test to the retest for each memory state (Bonferroni adjusted $p = .01$). Confidence increased significantly from initial test to retest for all states and was generally near ceiling on the retest: learning memory (87.9 vs. 96.1), $F(1,54) = 35.75$, $MSE = 51.66$, $\eta_p^2 = 0.40$, $p < .001$, related memory (84.0 vs. 98.9), $F(1,50) = 79.03$, $MSE = 71.88$, $\eta_p^2 = 0.61$, $p < .001$, feels familiar (57.3 vs. 77.2), $F(1,48) = 62.64$, $MSE = 155.41$, $\eta_p^2 = 0.57$, $p < 0.001$, and just know (92.0 vs. 99.4), $F(1,48) = 80.11$, $MSE = 16.59$, $p < .001$.

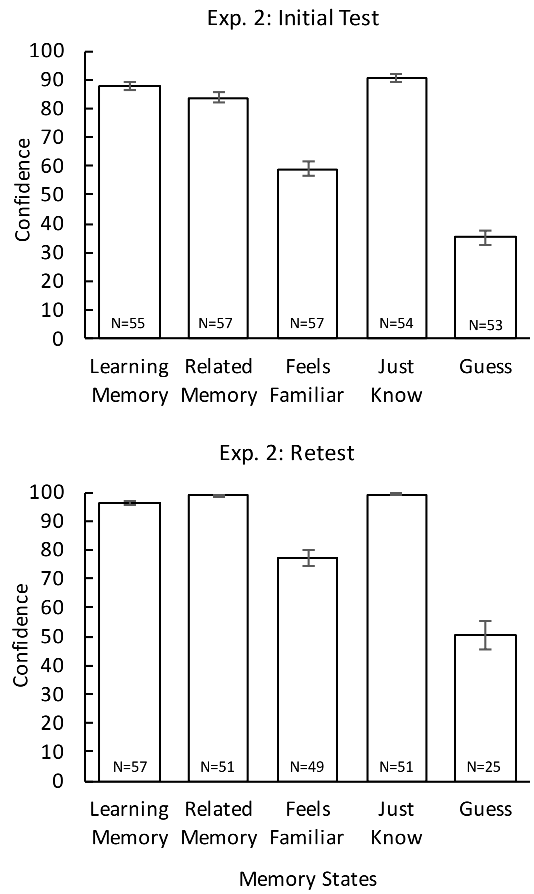

Figure 10. Mean confidence rating as a function of memory state across correct and incorrect answers in Experiment 2 initial test (top) and retest (bottom). Error bars indicate the standard error of the mean.

Hypercorrection

The retest was added in Experiment 2 in part to determine whether hypercorrection occurs when participants are asked to indicate their memory state during recall. To determine the level of hypercorrection, Goodman-Kruskal (1954) gamma correlations (a nonparametric measure that determines the relationship between ordinal variables; *G* = (concordant pairs-discordant pairs)/(concordant pairs+discordant pairs)) were calculated between the confidence ratings for incorrect answers on the recall portion of the initial test and accuracy for those same questions on the retest. Following Sitzman et al. (2015), to transform confidence into an ordinal variable, the scale was split to create ordinal categories: ratings of 0-24 on the confidence scale were assigned a value of 1 = *very low confidence*; ratings of 25-49 were assigned a value of 2 = *low confidence;* ratings of 50-74 were assigned a value of 3 = *high confidence*; and ratings of 75-100 were assigned a value of 4 = *very high confidence*. Figure 11 shows, for each confidence value, the proportion of incorrect answers on the initial test that were correct on the retest. To calculate gamma, concordant pairs were calculated by multiplying the number of very low confidence, incorrect answers on the initial test that were also correct on the retest by the sum of incorrect low, high, and very high confidence answers on the initial test that were correct on the retest. Discordant pairs were calculated by multiplying the number of very low confidence, incorrect answers that were correct on the retest by the sum of incorrect low, high, and very high confidence answers on the initial test that were also incorrect on the retest. Concordant and discordant pairs were calculated for each participant, and a gamma correlation was conducted for each participant from those values.

There was a moderate, positive relationship between confidence for incorrect answers and subsequent accuracy, $G = .41$, $t(25) = 2.62$, SE = 0.15, $p < .02$. Thus, a hypercorrection effect occurred such that high confidence incorrect initial test answers were associated with correct answers on the retest after feedback (e.g., Butler, Karpicke, & Roediger, 2008; Butterfield & Metcalfe, 2001, 2006; Fazio & Marsh, 2009; Sitzman et al., 2014; Sitzman et al., 2015).

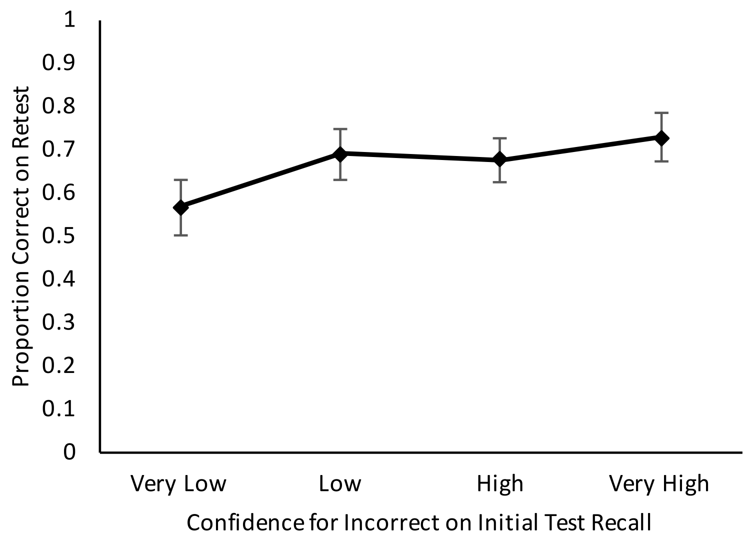

Figure 11. Proportion correct on the Experiment 2 retest as a function of confidence in incorrect answers on the initial test. Error bars indicate the standard error of the mean.

A gamma correlation could not be computed using the memory states due to their apparent non-ordinal relationship (i.e., numerical codes 1-5 cannot be assigned to the 5 qualitative states to reflect a known quantitative difference among them). Instead, I computed a hypercorrection-like index that parsed up the number of items assigned to each memory state on the retest as a function of whether those items were answered correctly, incorrectly, or were unanswered on the initial test. These partitions are presented in Figure 12 for both correct (left) and incorrect answers (right) on the retest. The number of incorrect answers was too low to analyze. The number of correct answers assigned to the learning memory state increased on the retest relative to the initial test (11.68 vs. 8.62), $F(1,109) = 20.36$, $MSE = 25.35$, $\eta^2 = 0.16$, $p < .001$. Thus, some correct answers on the initial test that had been assigned to other memory states shifted to the learning memory state on the retest, consistent with the possibility that the feedback provided access to previously inaccessible memories.

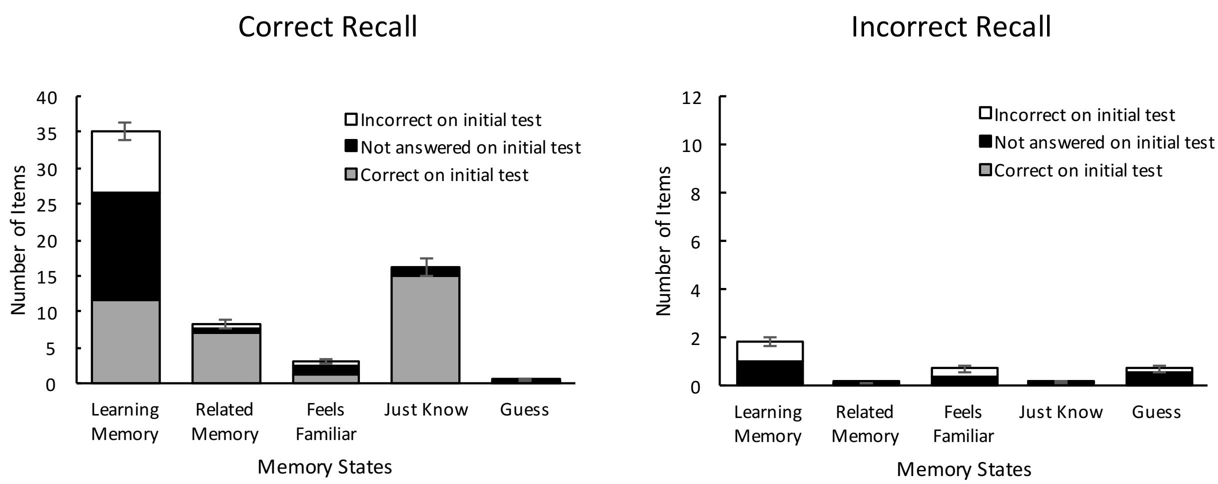

Figure 12. Bar heights indicate the number of items assigned to each memory state on the retest for correct answers (left) and incorrect answers (right). Bar partitions indicate the initial test outcome for each memory state on the retest. Error bars indicate standard error of the mean for the entire bar height.

The most striking result in the left panel of Figure 12, however, is that the bulk of the increase in learning memories on the retest came from those questions that were either incorrect and unanswered on the initial test. This result illustrates a novel version of a hypercorrection effect based on memory-state reports rather than confidence ratings. Learning memory gains in terms of correct answers on the retest were more frequent for unanswered questions on the initial test than for incorrect answers on the initial test (14.93 vs. 8.45), $F(1,109) = 48.52$, $MSE = 47.49$, $\eta^2 = 0.31$, $p < .001$. Thus, feedback improved later recall particularly well for the items that initially participants could not answer.

Chapter 4: General Discussion

My thesis presented a novel paradigm for exploring the memory states experienced while answering general-knowledge questions. Previous studies used other measures of metamemory with general-knowledge questions, such as confidence (Brewer & Sampaio, 2012; Butler et al., 2008; Butler, Fazio, & Marsh, 2011), feeling of knowing (Hart, 1965), and judgments of learning (Dunlosky & Nelson, 1992), but none assessed memory states in a recall task. In two experiments, participants answered a set of general-knowledge questions and identified their memory state after providing each answer (or their expected memory state for the answer when they could not provide one). Unlike the traditional dichotomous remember/know paradigm, participants were given 4 memory state options and 1 non-memory-state option, which offered them the potential to distinguish between two recollection states (learning memory vs. related memory), and between two non-recollection states (feels familiar vs. just know), as well as to be able to indicate guesses.

Experiment 1 measured the landscape of memory states reported for correct and incorrect answers, and for expected memory states, as well as the memory states experienced for the answers which were provided after each recall attempt. Four key novel findings emerged. First, participants commonly reported experiencing both of the recollection states, despite the general-knowledge task being traditionally used to tap semantic (vs. episodic) memory. Second, the two non-recollection memory states (feels familiar and just know) were dissociable in terms of use and accuracy, with the just know state being used more often and being much more accurate. Third, indeed, the just know state was as accurate as the two recollection-based states (which did not differ). And fourth, the memory states for the provided answers largely remained the same as the states initially reported for correct answers (i.e., they were stable in the face of feedback).

Experiment 2 replicated and extended Experiment 1. The design added four new features that yielded additional important findings. First and foremost, an immediate retest phase was added to determine how the distribution and accuracy of memory states changes after initial testing and answer feedback. The use and accuracy of the learning memory state increased after answer feedback, not surprisingly, but the key patterns among the memory states remained. Moreover, learning memories were prominent for items that participants answered incorrectly or failed to answer on the initial test. Second, rather than simply viewing each correct answer, participants received a multiple-choice question and indicated their memory state for their choice before being shown the answer. The multiple-choice accuracy pattern for the memory states matched the recall pattern. Third, I measured the source of each learning memory on the retest (i.e., feedback vs. pre-experiment). Although many of the additional correct answers that were experienced as learning memories on the retest were based on recollection of the feedback, there was also an increase in pre-experiment recollections. Fourth, a subset of participants provided confidence ratings after reporting their memory state. Confidence was found to be equally high for the just know state as for the two recollection states, and all three were much higher than the feels familiar state. Although the confidence difference implies that feels familiar and just know might differ only quantitatively, below I suggest that it could reflect an underlying qualitative difference that has been largely ignored in the remember/know literature. Additionally, a hypercorrection effect was also found such that higher (vs. lower) confidence incorrect answers on the initial test were associated with higher accuracy on the retest. A similar effect was found when examining performance on the initial test for learning memories. Thus, the experiments produced a number of novel findings to inform research on memory and metamemory for general-knowledge questions. Below I discuss my key findings and some of their implications in greater detail.

Drawing Distinctions Between Memory States for General Knowledge

Learning memory vs. related memory states. Of the hundreds of studies using the remember/know paradigm, very nearly all have begun with a study event, whether it be a list of words (e.g., Gardiner, 1988), a lecture (e.g., Conway et al., 1997), or a witnessed crime (e.g., Bodner & Richardson-Champion, 2007). As a result, the definitions of the recollection response participants are to use in these studies invariably refers to recollection of the study event. Experiencing recollection for test items from outside the study event is not relevant to the recognition decision participants must make. Rather, all that matters is whether they remember the items from the study event. In contrast, when piloting with general-knowledge questions, our lab group reported experiencing two different types of recollection. This led me to distinguish learning memories (recollection of where an answer was learned) from related memories (recollection of information related to the answers but not of learning the answer itself) in my experiments. This distinction was offered to participants on an exploratory basis, but my expectation was that learning memories (because they contain recollection of learning the answer itself) might be associated with higher accuracy (and with higher confidence in Experiment 2).

The similarities in the results for these two recollection memory states greatly outweighed the differences. Participants reported each state about equally often in Experiment 1 and on the initial test in Experiment 2, and the two states were associated with similar levels of accuracy and confidence. Moreover, both were similarly stable in the face of answer feedback. The only difference was that reports of learning memories for correct answers increased greatly on the retest in Experiment 2 (not surprisingly due to the answer feedback on the initial test) whereas reports of related memories actually decreased. This pattern also occurred for incorrect answers, though it was much less pronounced, suggesting that participants developed a response bias toward the learning memory state on the retest. Consistent with this possibility, accuracy was not greater for learning memories than related memories on the retest.

If the similarity in accuracy for learning and related forms of recollection holds up to further scrutiny, this result might be viewed as surprising. After all, it is intuitive that recollecting learning an answer should be more diagnostic of accuracy than recollecting related details that do not include recollecting learning an answer. On the other hand, the just know state was associated with similarly high accuracy. Therefore, perhaps recollection is overrated as a diagnostic tool for evaluating the accuracy of answers to general-knowledge questions. If having no recollection is as accurate as having learning memory recollection, then it seems reasonable that having related recollection is as accurate as having learning memory recollection.

Alternatively, some participants may not have understood or complied with the intended difference between the two recollection states. For example, one participant characterized their use of the learning memory state as “I chose learning memory mostly if I acquired that knowledge in school”, whereas related memory was chosen when they “… had a memory of it or learnt it through experience”. Thus, this participant defined learning memory more specifically as recollection of having learned the answer at school, whereas memories of learning the answer outside of school were deemed related memories, contrary to the provided definitions. However, other participants applied the intended distinction. For example, another participant’s example of a learning memory was “The question about the pancreas I remember learning in grade 12 bio with my teacher”, and their example of a related memory was “The question about the badmittion [sic] bird, I remembered it but thought of gym class first”. Although the raters were very good at distinguishing participants' definitions/examples/explanations for the two recollection states, their performance does not reify the distinction.

At present, the evidence for a qualitative difference between these two recollection states is weak, but further research may justify retaining this distinction. One approach worthy of exploration would be to provide a single recollection state option and ask participants to provide a written description of their recollection state for each recollection response. Raters could then classify these recollections into learning vs. related memory states, and analysis could determine whether the two states lead to different outcomes.

Feels familiar vs. just know states. Differences between the feels familiar and just know states were typically much greater than those between the learning and related memory states. More correct answers, and fewer incorrect answers, were associated with the just know state than the feels familiar state, thus accuracy was higher for the just know state (as was confidence in Experiment 2). As suggested above, these two states have often been used interchangeably in the metamemory literature, yet my results suggest that there are important quantitative, if not qualitative, differences underlying them. To take one example of the conflation of these memory states in previous research, Dewhurst and Anderson (1999) defined a know response as "…one in which you recognize the item because it feels familiar in this context, but you cannot recall its actual occurrence in the earlier phase of the experiment. You recognize the item purely on the basis of a feeling of familiarity" (p. 667). Conversely, a recent trend has seen researchers replacing the term *know* for *familiar* "... to indicate non-recollection, 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, Donaldson, MacKenzie, & Underhill, 1996). By conflating familiarity with knowing, researchers force participants to decide whether to categorize items associated with the familiarity state as “studied” or as “not studied” on a recognition test. As a result, reporting of knowing/familiarity might be more variable across participants, which could explain why less consistent results have been obtained for the effects of variables on non-recollective versus recollective memory (see Yonelinas, 2002, for a review).

Having shown that participants are willing and able to distinguish familiarity from knowing, the question remains as to whether these two non-recollective states differ qualitatively or only quantitatively. Is knowing simply a stronger memory state than familiarity, as implied by Donaldson et al.’s (1996) decision to replace the term “knowing” for the term “familiarity”? My results are consistent with this claim, given that the just know state was associated with higher confidence ratings than the feels familiar state. Participants may use differences in confidence to assign the feels familiar vs. just know memory state, rather than basing this distinction on whether they experience a feeling of familiarity or a just know. In turn, the assignment of higher confidence answers to the just know state this would explain why it led to greater accuracy than the feels familiar state, given that accuracy and confidence in recall are correlated (Brewer & Sampaio, 2012; Koriat, 2008). To take another example, in Experiment 1 the just know state was more stable than the feels familiar state in the face of answer feedback (Figure 2). Correct initial answers that felt familiar shifted similarly toward a learning memory, related memory, or just know state for the provided answer. In contrast, initial answers assigned to the just know state typically remained in that state. This outcome is also consistent with just knowing reflecting higher confidence.

Raters were highly accurate at classifying the descriptions/examples participants provided for the feels familiar vs. just know states. Although this could be taken as evidence that participants treated these states as qualitatively different, it is also possible that the raters differentiated these states by participants' verbiage that indicated confidence. For example, one participant defined the just know state as "… when you are positively [sic] that it is a specific answer but you don't remember where you got that info from", whereas the feels familiar state was when "you are not sure where you heard the answer or if it is correct, but you have a feeling about it". Although confidence terms were avoided in the provided definitions, participants often included confidence words like "sure" and "positive" in their responses. Raters could have used this language to accurately classify participants' responses.

On the other hand, there is a clear difference in the subjective experience of an answer that "feels familiar" and an answer that one "just knows" that should not be ignored simply because this difference is associated with a difference in confidence ratings. The nagging feeling in familiarity is noticeably absent from just knowing, and was likely used by at least a subset of participants to differentiate between the two states. Critically, given that confidence was similarly high for the learning memory, related memory, and just know states (see Figure 10), participants must have relied on a basis other than differences in memory strength to partition their responses into these three categories. That basis would be the subjective differences in the memory states laid out in the experiment. It is therefore not a great leap to suggest that participants attempted to classify their non-recollection states into just know and feels familiar categories on the basis of their subjective experiences as well. Of course, this claim will require further scrutiny and testing. Regardless, taking into account the magnitude of the differences in the use and performance of the two non-recollection states, echoing Williams and Moulin (2014), my experiments lead me to strongly recommend that researchers distinguish between these two memory states when subjective metamemory states are to be measured.

Recollection vs. non-recollection. My thesis findings also inform the larger debate about whether recollection and non-recollection (whether labeled as "familiarity" or as "knowing") reflect qualitatively different processes—as originally argued by Tulving (1985) and Gardiner (1988), and more recently by others including Yonelinas (2002)—or a single process that yields memories that differ merely quantitatively in their strength—as originally argued by Hirshman (1995) and Donaldson (1996), and more recently by others including Dunn (2008). In my study, just knowing was very similar to both recollection states in terms of accuracy and critically, also in terms of confidence. Given that confidence is an indicator of memory strength, clearly participants are distinguishing these memory states on a basis other than memory strength. My finding provides a striking new challenge to the claim that remembering and *knowing* merely differ quantitatively. Instead, only the feels familiar state differed from all three of these states and could thus indicate a memory state whose use is based on memory strength. Of course, given that previous studies have often conflated knowing and familiarity (as discussed above), a quantitative difference between recollection and non-recollection would have been supported. However, when knowing and familiarity are clearly separated, as was the case in my study, then there is no evidence that recollection and *knowing* reflect different points on a continuum of memory strength.

In turn, this pattern begs the question of what the just know state and the recollection states have in common if they cannot be differentiated in terms of accuracy and confidence. Having a recollection experience was expected to signal accuracy, but surprisingly, the absence of a recollection experience—so long as it was substituted by a just know experience—was equally good at signaling accuracy. To understand this pattern, it will be helpful for future research to seek to identify the key difference that underlies the just knowing vs. feels familiarity state. The definition of feels familiar referred to a nagging feeling of familiarity for the answer, while just knowing was described as the absence of a feeling state for the answer. This absence of a nagging feeling might instead be characterized as the presence of a feeling of certainty, in line with the high confidence ratings given for just know responses. Thus, a feeling of certainty may be the key common feature that links the recollection and just knowing states and that in turn differentiates all three states from the feels familiar state. Such a possibility is worth exploring in future research.

Erasing a Distinction Between Episodic and Semantic Memory?

Underlying the preceding debate about the distinction captured by recollective versus non-recollective memory states was the striking prevalence of episodic memory (i.e., learning and related memory states) for answers to general-knowledge questions in the present study. The task of answering general-knowledge questions has been used as a hallmark method of measuring implicit conceptual or semantic memory (e.g., see Blaxton, 1989). If this task taps semantic memory, then why was episodic recollection so frequently experienced? One response would be to argue that the answers were drawn from semantic memory, but sometimes also resulted in retrieval of related recollections from episodic memory (e.g., Benjamin, Bjork, & Schwartz, 1998). However, this possibility would be difficult to distinguish from the simpler claim that the answers leading to recollective states were drawn from episodic memory.

Indeed, the claim that episodic memory, whether or not it yields recollection, is the basis of all memory retrieval is a hallmark of instance and exemplar based theories of memory (e.g., Franklin & Mewhort, 2015; Hintzman, 1984; Logan, 1988; 2002; Whittlesea, 1997). From this perspective, the general-knowledge test, as with any memory test, taps a memory store of prior instances/experiences; by such accounts semantic memory does not exist as a separate memory store. My results fit nicely with such theories, so long as they allow for the emergence of qualitatively different subjective memory states for the products of memory retrieval attempts.

The similarity between recollection states and just knowing is problematic for Tulving's (1985) claim that these states reflect access to episodic and semantic memory systems, respectively. However, as reviewed above, the feels familiar and just know states differed, thus it may be possible to retain a distinction between episodic and semantic memory by arguing that the feels familiar state reflects retrieval from semantic memory. By this tack, the just know state would reflect retrieval from the episodic memory system *even though it does not give rise to any recollection of episodic detail*. Such a possibility would be completely contrary to the traditional definition of episodic memory, but would fit with instance/exemplar based memory models. At a minimum, the claim that the general-knowledge task only taps semantic memory can be ruled out on the basis of my findings. However, my findings contribute to a broader challenge to the episodic/semantic memory distinction put up by others (e.g., Blaxton, 1989; Jacoby, 1983; Roediger, 1990).

Of course, the episodic/semantic distinction (or more neutrally, the autonoetic/noetic distinction) is supported by many other forms of evidence (see Yonelinas, 2002). For example, amnesic patients often show impaired access to episodic/recollective memory and autonoetic consciousness but normal access to semantic/familiarity memory and noetic consciousness (e.g., Graf, Squire, Mandler, 1984; Tulving, 1985; Warrington & Weiskrantz, 1970). My findings do not directly challenge these other forms of evidence for the episodic/semantic distinction. Nonetheless, my findings do invite further consideration of how best to carve the nature of long-term memory at its joints, and even whether it warrants carving up at all. To be sure, the general-knowledge test does not provide a pure measure of semantic long-term memory.

The Source of Learning Memories During a Retest

Participants in Experiment 2 reported far more learning memories on the retest than on the initial test (see Figure 2). Their recollection source judgments revealed that their learning memories for correct answers given on the retest were based on having seen the answers during the initial test (i.e., feedback was often reported as the basis of their learning memories). However, the number of correct answers that were reported as having been based on a pre-experiment learning memory also increased on the retest. Thus, viewing the answer feedback on the initial test appears to have allowed participants to gain access to pre-experiment recollections for additional items. This finding nicely captures an experience that commonly occurs while playing trivia games, wherein after failing to come up with a correct answer and then learning the answer, one is then able to recollect details about it.

Importantly, the salience of memory for the feedback during the retest did not eliminate pre-experiment recollections for correctly answered items. As shown in Figure 9, the rate of learning memory states reported for correct initial answers went up on the retest, with further increases in the rate of learning memory states occurring primarily for items that were answered incorrectly or that were not answered on the initial test. Thus, pre-experiment recollections were not overshadowed by the recency and salience of the feedback on the immediate retest. This could indicate the greater distinctiveness of pre-experiment recollections relative to feedback recollections. Pre-experiment memories would have been formed in different places at different times, whereas the feedback memories would likely be more uniform (i.e., recollecting having seen the answer on the screen).

The feedback appears to have been particularly salient during the retest for items the participant either could not answer or answered incorrectly on the initial test. Incorrect and unanswered questions often yielded learning memories on the retest (see Figure 12), which I consider to be a novel type of hypercorrection effect (Butterfield & Metcalfe, 2001; Fazio & Marsh, 2009), which refers to a pattern in which high confidence incorrect answers on an initial test are more likely to be correct on a retest. The increase in learning memories for incorrect/unanswered questions corroborates the hypercorrection literature by showing that surprising feedback often leads to learning/correction, and adds to it by revealing that this learning/correction often yields a recollection experience. The more traditional hypercorrection effect (i.e., based on confidence ratings during the initial test) was found in Experiment 2, and was comparable in strength to other reports of hypercorrection for general-knowledge questions following answer feedback (e.g., Butler et al., 2008; Butterfield & Metcalfe, 2001; Cyr & Anderson, 2013; Fazio & Marsh, 2009). Thus, both confidence and response-type hypercorrection effects can both be obtained when memory states are assessed in a recall task.

Limitations and Future Directions

In addition to the areas for future research identified above, some limitations of the present work invite additional follow-ups. One such limitation was that the relatively high accuracy rates (relative to what was anticipated from the norms of Tauber et al., 2013) resulted in fewer incorrect and expected memory state observations than anticipated. I was therefore unable to examine in detail the memory states for incorrect answers and unanswered questions. Replicating my study using more difficult items (or even by selecting deceptive questions, see Brewer & Sampaio, 2012) would allow more detailed analysis of such memory states.

Another limitation was that the immediate retest did not yield much change in the memory states, save for the increase in learning memories due to the answer feedback. The immediate retest allowed some research questions to be answered, such as whether memory for salient feedback overshadows pre-experiment memory. However, it also prevented an investigation of how memory states for general knowledge (and for answer feedback) shift over time. Future research could institute a delay before the retest to explore the potential for such shifts and whether accuracy and confidence shift in kind. For example, if the recollection and just know states both reflect access to episodic memory, then they should not shift to a feels familiar state over a longer delay, whereas the feels familiar state would reflect access to semantic memory and so should not shift to either a recollection or just know state. Of course, there is evidence that knowledge learned in lectures often shifts from “remembered” to “known”, thus the same may be true for general-knowledge learned long ago (e.g., Conway et al., 1997).

Finally, it may be useful to examine individual differences and material-based factors. For example, older adults might experience less recollection than younger adults (Morson et al., 2015; Souchay et al., 2007). Additionally, it would be interesting to explore how the frequency and distribution of memory states differ for trivia experts relative to novices. It would also be interesting to examine whether different categories of general knowledge (e.g., sports, history, etc.) elicit differing memory state distributions.

Implications and Conclusions

The present results have implications for educational contexts. My thesis suggests that memory for general knowledge, including information learned decades earlier in school, often remains episodic/recollective in nature. If different memory states are found to be differentially linked to accuracy, then learners could be trained to evaluate their memory state carefully during the learning process, as well as during their application of knowledge on tests and other forms of evaluation. Additionally, the present study adds to growing literature on the benefits of testing and feedback (e.g., Roediger & Karpicke, 2006; Fazio & Marsh, 2009); performance improved on an immediate retest following an initial test and feedback (the effects of which could be separated in future studies). Whether these learning effects, including memory for feedback, are retained over time also remains to be explored.

In summary, my thesis provided the first in-depth exploration of the prevalence and accuracy of various memory states for general-knowledge questions. It provides some challenges for the traditional distinction between remembering and knowing, but it also provides some initial evidence that it could be useful to distinguish different recollective states (learning vs. related recollections) and that it is critical to distinguish different non-recollective states (feels familiar vs. just know). Taken together, my findings initiate a call to action for researchers in this field to further explore the metamemory states associated with memory retrieval. Such explorations have the potential to reveal essential distinctions in long-term memory that could assist in creating better strategies for helping people assess their learning, and perhaps even for retrieving information from memory.

References

Allwood, C. M. & Bjorhag, C. G. (2008). Mood and realism of confidence judgements of one's own answers to general-knowledge questions. *Scandinavian Journal of Psychology, 32*, 358-371. doi:10.1111/j.1467-9450.1991.tb00887.x

Altman, D. G. (1991). Practical statistics for medical research. London: Chapman and Hall.

Aue, W. R., Criss, A. H., & Prince, M. A. (2015). Dynamic memory searches: Selective output interference for the memory of facts. *Psychonomic Bulletin & Review, 22*, 1798-1806. doi:10.3758/s13423-015-0840-5

Backman, L. & Nilsson, L. G. (1996). Semantic memory functioning across the adult life span. *European Psychologist, 1*, 27-33. doi:10.1027/1016-9040.1.1.27

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. doi:10.1080/09658210802360603

Benjamin, A. S., Bjork, R. A., & Swartz, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General, 127*, 55-68. doi:10.1037/0096-3445.127.1.55

Blaxton, T.A. (1989). Investigating dissociations among memory measures: Support for a transfer-appropriate processing framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 657-668. doi:10.1037/0278-7393.15.4.657

Bodner, G. E., & Richardson-Champion, D. D. L. (2007). Remembering is in the details: Effects of test-list context on memory for an event. *Memory, 15*, 718-729. doi:10.1080/09658210701674682

Bornstein, B. H., & Zickafoose, D. J. (1999). “I know I know it, I know I saw it”: The stability of the confidence-accuracy relationship across domains. *Journal of Experimental Psychology: Applied, 5*, 76-88. doi:10.1037/1076-898X.5.1.76

Brennan, T., Baguley, T., Bright, J., & Bruce, V. (1990). Resolving semantically induced tip-of-the-tongue states for proper nouns. *Memory & Cognition, 18*, 339-347. doi:10.3758/BF03197123

Brewer, W. F., & Sampaio, C. (2012). The metamemory approach to confidence: A test using semantic memory. *Journal of Memory and Language, 67*, 59-77. doi:10.1016/j.jml.2012.04.002

Brown, A. S. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin, 109*, 204-223. doi:10.1037/0033-2909.109.2.204

Brown, R., & McNeill, D. (1966). The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior, 5*, 325-337.

Butler, A. C., Karpicke, J. D., & Roediger, H. L., III. (2008). Correcting a metacognitive error: Feedback increases retention of low-confidence correct responses. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 918-928. doi:10.1037/0278-7393.34.4.918

Butler, A. C., Fazio, L. K., & Marsh, E. J. (2011). The hypercorrection effect persists over a week, but high-confidence errors return. *Psychonomic Bulletin and Review, 18*, 1238-1244. doi:10.3758/213423-011-0173-y

Butterfield, B., & Metcalfe, J. (2001). Errors committed with high confidence are hypercorrected. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 1491-1494. doi:10.1037//0278-7393.27.6.1491

Butterfield, B., & Metcalfe, J. (2006). The correction of errors committed with high confidence.

Metacognition Learning, 1, 69-84. doi:10.1007/s11409-006-6894-z

Buzztime. (2016, May 6). NTN Buzztime, Inc. reports first-quarter 2016 results [Press release]. Retrieved from http://www.buzztime.com/business/blog/ntn-buzztime-inc-reports-first-quarter-2016-results/

Chilton, M. (2014). Trivial Pursuit: 10 facts about the game. *The Telegraph*. Retrieved from http://www.telegraph.co.uk/culture/culturenews/11293880/Trivial-Pursuit-10-facts-about-the-game.html

Chua, E. F. & Ahmed, R. (2016). Electrical stimulation of the dorsolateral prefrontal cortex improves memory monitoring. *Neuropsychologia, 85*, 74-79. doi:10.1016/j.neuropsychologia.2016.03.008

Cyr, A. A. & Anderson, N. D. (2013). Updating misconceptions: Effects of age and confidence. *Psychonomic Bulletin & Review, 20*, 574-580. doi:10.3758/s13423-012-0357-0

Cleary, A. M., & Claxton, A. B. (2015). The tip-of-the-tongue heuristic: How tip-of-the-tongue states confer perceptibility on inaccessible words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*, 1533-1539.

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. doi:10.1037/0096-3445.126.4.393

Dewhurst, S. A., & Anderson, S. J. (1999). Effects of exact and category repetition in true and false recognition memory. *Memory & Cognition, 27*, 665-673. doi:10.3758/BF03211560

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.

doi:10.1002/acp.1517

Donaldson, W. (1996). The role of decision processes in remembering and knowing. *Memory & Cognition, 24*, 523–533. doi:10.3758/BF03200940

Donaldson, W., MacKenzie, T. M., & Underhill, C. F. (1996). A comparison of recollective memory and source monitoring. *Psychonomic Bulletin & Review, 3*, 486-490. doi:10.3758/BF03214551

Dudukovic, N. M., & Knowlton, B. J. (2006). Remember-know judgments and retrieval of contextual details. *Acta Psychologica*, *122*, 160-173. doi:10.1016/j.actpsy.2005.11.002

Dunlosky, J., Mueller, M. L., & Thiede, K. W. (2016). Methodology for investigating human metamemory: Problems and Pitfalls. In J. Dunlosky & S. K. Tauber (Eds.), *The Oxford Handbook of Metamemory* (23-38). New York, NY: Oxford University Press.

Dunlosky, J., & Nelson, T. O. (1992). Importance of the kind of cue for judgments of learning (JOL) and the delayed-JOL effect. *Memory & Cognition, 20*, 374-380. doi:10.3758/BF03210921

Dunn, J. C. (2004). Remember-know: A matter of confidence. *Psychological Review, 111*, 524-542. doi:10.1037/0033-295X.111.2.524

Dunn, J. C. (2008). The dimensionality of the remember-know task: A state-trace analysis. *Psychological Review, 115*, 426-446. doi:10.1037/0033-295X.115.2.426

Fazio, L. K., & Marsh, E. J. (2009). Surprising feedback improves later memory. *Psychonomic Bulletin & Review, 16*, 88-92. doi:10.3758/PBR.16.1.88

Ferguson, A. M., McLean, D., & Risko, E. F. (2015). Answers at your fingertips: Access to the Internet influences willingness to answer questions. *Consciousness and Cognition, 37*, 91-102. doi:10.1016/j.concog.2015.08.008

Fischhoff, B. (1975). Hindsight ≠ foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance, 1*, 288-299. doi:10.1037/0096-1523.1.3.288

Flavell, J. H., & Wellman, H. M. (1977). Metamemory. In R. V. Kail & J. W. Hagen (Eds.), *Perspectives on the development of memory and cognition* (p. 3-33). Hillsdale, NJ: Erlbaum.

Franklin, D. R. J., & Mewhort, D. J. K. (2015). Memory as a hologram: An analysis of learning and recall. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 69*, 115-135. doi:10.1037/cep0000035

Furnham, A. & Chamorro-Premuzic, T. (2006). Personality, intelligence and general knowledge. *Learning and Individual Differences, 16*, 79-90. doi:10.1016/j.lindif.2005.07.002

Gardiner, J. M. (1988). Functional aspects of recollective experience. *Memory & Cognition, 16*, 309-313. doi:10.3758/BF03197041

Gardiner, J. M., Ramponi, C., & Richardson-Klavehn, A. (1998). Experiences of remembering, knowing, and guessing. *Consciousness and Cognition, 7*, 1-26. doi:10.1006/ccog.1997.0321

Goodman, L. A., & Kruskal, W. H. (1954). Measures of association for cross classifications. *Journal of the American Statistical Association, 49*, 732-764. doi:10.1007/978-1-4612-9995-0_1

Graf, P., Squire, L. R., & Mandler, G. (1984). The information that amnesic patients do not forget. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 164-178. doi:10.1037/0278-7393.10.1.164

Hart, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational*

Psychology, 56, 208-216. doi:10.1037/h0022263

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. doi:10.1002/acp.947

Hicks, J. L., & Marsh, R. L. (2002). On predicting the future states of awareness for recognition of unrecallable items. *Memory & Cognition, 30*, 60-66. doi:10.3758/BF03195265

Hintzman, D. L. (1984). Episodic versus semantic memory: A distinction whose time has come – and gone? The Behavioral and Brain Sciences, 7, 223-268. doi:10.1017/S0140525X00044435

Hirshman, E. (1995). Decision processes in recognition memory: Criterion shifts and the list-strength paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 302-313. doi:10.1037/0278-7393.21.2.302

Jacoby, L. L. (1983). Remembering the data: Analyzing interactive processes in reading. *Journal of Verbal Learning & Verbal Behavior, 22*, 485-508. doi:10.1016/S0022-5371(83)90301-8

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

Koriat, A. (1995). Dissociating knowing and the feeling of knowing: Further evidence for the accessibility model. *Journal of Experimental Psychology: General, 124*, 311-333. doi:10.1037/0096-3445.124.3.311

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

Laham, D., & Steinhart, D. (1998). Latent Semantic Analysis @ CU Boulder. Retrieved from http://lsa.colorado.edu

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95*, 492-527. doi:10.1037/0033-295X.95.4.492

Logan, G. D. (2002). An instance theory of attention and memory. *Psychological Review, 109*, 376-400. doi:10.1037/0033-295X.109.2.376

McCabe, D. P., & Soderstrom, N. C. (2011). Recollection-based prospective metamemory judgements are more accurate than those based on confidence: Judgements of remembering and knowing (JORKs). *Journal of Experimental Psychology: General*, *140*, 605-621. doi:10.1037/a0024014

Metcalfe, J., Schwartz, B. L., & Joaquim, S. G. (1993). The cue familiarity heuristic in metacognition. *Journal of Experimental Psychology: Learning, Memory, & 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. doi:10.1016/j.jml.2013.01.001

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

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin, 95*, 109-133. doi:10.1037/0033-2909.95.1.109

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *The Psychology of Learning and Motivation, 26*, 125-173. doi:10.1016/S0079-

7421(08)60053-5

Princeton University. (2010). WordNet. Retrieved from https://wordnet.princeton.edu/

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

Roediger, H. L. (1990). Implicit memory: Retention without remembering. *American Psychologist, 45*, 1043-1056. doi:10.1037/0003-066X.45.9.1043

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*, 249-255. doi:10.1111/j.1467-9280.2006.01693.x

Schwartz, B. L. (2001). The relation of tip-of-the-tongue states and retrieval time. *Memory & Cognition, 29*, 117-126. doi:10.3758/BF03195746

Schwartz, B. L. (2008). Working memory load differentially affects tip-of-the-tongue states and feeling-of-knowing judgment. *Memory & Cognition, 36*, 9-19. doi:10.3758.MC.36.1.9

Sitzman, D. M., Rhodes, M. G., & Tauber, S. K. (2014). Prior knowledge is more predictive of error correction than subjective confidence. *Memory & Cognition, 42*, 84-96. doi:10.3758/s13421-013-0344-3

Sitzman, D. M., Rhodes, M. G., Tauber, S. K., & Liceralde, V. R. T. (2015). The role of prior knowledge in error correction for younger and older adults. *Aging, Neuropsychology, and Cognition, 22*, 502-516. doi:10.1080/13825585.2014.993302

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.

doi:10.1016/j.concog.2006.11.002

Tauber, S. K., Dunlosky, J., Rawson, K. A., Rhodes, M. G., & Sitzman, D. M. (2013). General knowledge norms: Updated and expanded from the Nelson and Narens (1980) norms. *Behavior Research Methods, 45*, 1115-1143. doi:10.3758/s13428-012-0307-9

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381–403). New York: Academic Press.

Tulving, E. (1985). Memory and consciousness. Canadian Psychology, 26, 1–12. doi:10.1037/h0080017

Warrington, E. K., & Weiskrantz, L. (1974). The effect of prior learning on subsequent retention in amnestic patients. *Neuropsychologia, 12*, 419-428. doi:10.1016/0028-3932(74)90072-4

Williams, H. L., & Moulin, C. J. A. (2014). Know versus familiar: Differentiating states of awareness in others' subjective reports of recognition. *Memory, 23*, 981-990. doi:10.1080/09658211.2014.945460

Whittlesea, B. W. A. (1997). Production, evaluation, and preservation of experiences: Constructive processing in remembering and performance tasks. In D. L. Medin (Ed.), *The psychology of learning and motivation: Advances in research and theory,* Vol. 37, pp. 211-264). San Diego, CA: Academic Press.

Wixted, J. T., & Mickes, L. (2010). A continuous dual-process model of remember/know judgments. *Psychological Review, 11*7, 1025-1054. doi:10.1037/a0020874

Wojcik, D. Z., Moulin, C. J. A., & Souchay, C. (2013). Metamemory in children with autism: Exploring "feeling-of-knowing" in episodic and semantic memory. *Neuropsychology, 27*, 19-27. doi:10.1037/a0030526

Yaniv, I., & Meyer, D. E. (1987). Activation and metacognition of inaccessible information:

Potential bases for the incubation effect in problem solving. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 13*, 187-205.

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. doi:10.1006/jmla.2002.2864