REMEMBERING AND KNOWING: EXPLORING SUBJECTIVE REPORT, FAMILIARITY, AND CONFIDENCE

Helen Louise Williams

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds
Institute of Psychological Sciences

June 2011

ACKNOWLEDGMENTS

Firstly I would like to thank my supervisor, Chris Moulin, for his guidance, sense of humour, many hours of debate, and for such moments as sitting on the floor of his office classifying justification statements as Know or Familiar. Chris has inspired me to be excited about the research process since the final year of my Undergraduate degree and has become a great friend along the way. I am also indebted to my second supervisor, Martin Conway, for his advice and the opportunity of working as his research assistant, which allowed me to start to develop as a researcher prior to my PhD.

I am also indebted to Céline Souchay, Matti Laine, and Jane Berry for providing invaluable professional and personal advice and broadening my psychological knowledge base. Extreme gratitude also goes to all members of MoulinLab© and SouchayLab©, for providing novel ideas, interesting papers, wonderful cakes, and exciting scientific discussion every Friday.

Big thanks must also go to Sarah Smith, Christine Wells, and Clare Rathbone. Each has provided their own unique kind of support, advice, humour, and friendship over the last few years, and I can't imagine what the PhD would have been like without them. Other friends in Leeds, Hertfordshire, Richmond, and around the world have also provided great fun and distraction from work. Immense thanks also goes to my family: mum Ingrid, dad John, sister Nicola, and all the animals. They have given me tremendous support during the PhD and going home has always been a lovely diversion from experiments and writing.

This thesis is dedicated to two women. Firstly my mum, Ingrid Williams, who inspired me towards academic research since I sat through her practice conference slideshows from an early age. Mum has always been a voice of reason for any academic anxiety and has provided a wealth of advice throughout my PhD. It is also dedicated to my grandmother, Milda Weinberg (1913–2007), who participated in my first Undergraduate case-study project. The memories she retrieved to cue words got me interested in the distinction, or lack thereof, between episodic and semantic memory.

ABSTRACT

This thesis examined how people make and understand judgments of subjective experience using the categories of Remember, Know, Familiar, and Guess (R, K, F, G), and represents the first attempt to use all four categories in a standard episodic recognition task. The key findings of this body of work are that Know and Familiar categories of subjective experience can be reliably differentiated, as can each of the four subjective experience categories from confidence. Chapter 2 examined lay understanding of subjective experience by asking participants to examine others' memory justification statements. Participants reliably differentiated the four justification types (R, K, F, G) in terms of confidence (Experiment 2.1) and subjective experience (Experiments 2.2 and 2.3); manipulations of confidence influenced assignment of justifications to subjective experience categories for some types of justification more than others (Experiment 2.2); and participants were able to divide justification statements into Know and Familiar when no definitions of those concepts were provided (Experiment 2.4). Chapter 3 investigated the influence of experimentally imposed familiarity on subjective experience. Pre-exposure of target and lure items led to impaired recognition across all experiments, but differences in subjective experience were only observed when pre-exposure was performed between-subjects (Experiment 3.1 vs. 3.2 and 3.3). Participants were able to use recollection strategically to overcome the familiarity induced by pre-exposure. Chapter 4 compared source, confidence, and subjective experience judgments and demonstrated subjective experience judgments to be more sensitive to source accuracy than confidence judgments; confidence judgments were more lenient than subjective experience; and confidence judgments were more affected by source manipulations than were subjective experience judgments. Across the thesis, analysis of reaction time also demonstrated reliable differences between and within judgment types. This thesis found critical differences between Remember, Know, and Familiar. Know and Familiar judgments were shown to dissociate on recognition accuracy, source accuracy, confidence, and response time. In contrast, Remember and Know judgments were only shown to be differentiated by source accuracy. The findings have implications for methodological and theory development and are discussed in terms of single- and dual-process accounts of memory.

TABLE OF CONTENTS

ACKNOWLEDGMENTS III

ABSTRACT IV

TABLE OF CONTENTS V

LIST OF FIGURES X

LIST OF TABLES XIV

ABBREVIATIONS XVI

1. GENERAL INTRODUCTION 17

1.1. OVERVIEW 17

1.2. RECOLLECTION AND FAMILIARITY; REMEMBERING AND KNOWING 18

1.3. IMPLICATIONS AND APPLICATIONS 19

1.4. THEORETICAL ISSUES 21

1.4.1. Dual-process versus single-process accounts of recognition memory 21

1.4.2. Dual-process models 21

1.4.3. Single-process models 24

1.4.4. Single-process accounts of Remember-Know data 26

1.4.5. Recent additions to the theoretical debate 28

1.5. MEASURING RECOLLECTION AND FAMILIARITY 30

1.5.1. Task-dissociation methods 30

1.5.2. Process-estimate methods 34

1.6. THE REMEMBER-KNOW PARADIGM 38

1.6.1. Experimental dissociations 38

1.6.2. Special populations 39

1.6.3. Evidence from neuropsychology 41

1.6.4. Comparisons with confidence judgments 45

1.6.5. Measurement issues within the Remember-Know paradigm 47

1.7. AIMS OF THIS THESIS 56

2. USING LAY CONCEPTIONS OF SUBJECTIVE EXPERIENCE TO EXPLORE REMEMBERING AND KNOWING, AND FINDING FAMILIAR 59

2.1. GENERAL INTRODUCTION 59

2.1.1. The influence of terminology and confidence on Remember-Know judgments 61
2.1.2. Participants' understanding of Remember-Know definitions ... 62
2.1.3. Aims of the current experiments ... 66
2.1.4. Splitting Know into Know and Familiar ... 67
2.2. Experiment 2.1: Lay understanding of subjective experiences in terms of confidence ... 68
2.2.1. Introduction ... 68
2.2.2. Method ... 69
2.2.3. Results ... 70
2.2.4. Discussion ... 73
2.3. Experiment 2.2: Lay understanding of subjective experience in terms of Remember, Know, Familiar, and Guess judgments ... 75
2.3.1. Introduction ... 75
2.3.2. Method ... 76
2.3.3. Results ... 78
2.3.4. Discussion ... 87
2.4. Experiment 2.3: Lay understanding of subjective experience when confidence is manipulated ... 91
2.4.1. Introduction ... 91
2.4.2. Method ... 92
2.4.3. Results ... 94
2.4.4. Discussion ... 102
2.5. Experiment 2.4: Lay-person categorisation of Know justifications into Know and Familiar ... 105
2.5.1. Introduction ... 105
2.5.2. Method ... 106
2.5.3. Results ... 107
2.5.4. Discussion ... 114
2.6. General discussion ... 117
2.6.1. The split of Know into Know and Familiar ... 117
2.6.2. Confidence and subjective experience ... 118
3. The Use of Recollection When Items are Familiar ... 121
3.1. General introduction ... 121
3.1.1. Word frequency effects ... 122
3.1.2. Word frequency effects and subjective experience ... 122

3.1.3. Experimental manipulations of familiarity 124
3.1.4. Experimental manipulations of familiarity with measures of subjective experience 128
3.1.5. Experimental design for subjective experience experiments 132
3.1.6. Aims of the current experiments 133
3.2. Experiment 3.1: Recollection in associative recognition 134
3.2.1. Introduction 134
3.2.2. Method 135
3.2.3. Results 137
3.2.4. Discussion 143
3.3. Experiment 3.2: Recollection when familiarity is manipulated within-subjects 145
3.3.1. Introduction 145
3.3.2. Method 145
3.3.3. Results 147
3.3.4. Discussion 151
3.4. Experiment 3.3: Recollection in item and non-associative recognition in a blocked within-subjects design 153
3.4.1. Introduction 153
3.4.2. Method 155
3.4.3. Results 158
3.4.4. Discussion 169
3.5. General Discussion 174
3.5.1. Memory performance 174
3.5.2. Subjective experience 175
3.5.3. Reaction time 179
3.5.4. Experimental design for subjective experience experiments 180
3.5.5. Conclusions 181
4. Subjective Experience, Confidence, and Source: Differences, Influence, and Reaction Time 183
4.1. General introduction 183
4.1.1. Comparing subjective experience and confidence judgments 184
4.1.2. The influence of one judgment on another judgment 186
4.1.3. Reaction time data associated with confidence and subjective experience judgments 187

4.2. EXPERIMENT 4.1: SUBJECTIVE EXPERIENCE AND CONFIDENCE JUDGMENTS 190
4.2.1. Introduction 190
4.2.2. Method 191
4.2.3. Results 194
4.2.4. Discussion 209
4.3. EXPERIMENTS 4.2 AND 4.3: SOURCE JUDGMENTS, SUBJECTIVE EXPERIENCE, AND CONFIDENCE . 212
4.3.1. Comparing subjective experience and source judgments 212
4.3.2. Comparing source and confidence judgments 213
4.3.3. The influence of one judgment on another 214
4.3.4. Reaction time for judgments of source 216
4.3.5. The current experiments 217
4.3.6. Methods 218
4.3.7. Results 219
4.3.8. Discussion 238
4.4. GENERAL DISCUSSION 242
4.4.1. Comparing subjective experience and confidence 243
4.4.2. Comparing source, subjective experience, and confidence 244
4.4.3. Reaction time for recognition 245
4.4.4. Reaction time to make judgments 246
4.4.5. Concluding remarks 247
5. GENERAL DISCUSSION 248
5.1. OVERVIEW 248
5.2. SUMMARY OF FINDINGS 248
5.2.1. Chapter 2 249
5.2.2. Chapter 3 252
5.2.3. Chapter 4 253
5.2.4. Reaction time for recognition decisions 256
5.2.5. Reaction time to make judgments 259
5.3. THE SEPARATION OF KNOW AND FAMILIAR 262
5.4. THE DISSOCIATION OF CONFIDENCE AND SUBJECTIVE EXPERIENCE 266
5.5. HOW DO OBJECTIVE MANIPULATIONS AFFECT SUBJECTIVE EXPERIENCE? 270
5.6. METHODOLOGY IN SUBJECTIVE EXPERIENCE RESEARCH 272
5.7. APPLICATIONS AND FUTURE RESEARCH 273
5.8. CONCLUSIONS 276

REFERENCES **277**

APPENDIX A: REMEMBER, KNOW, FAMILIAR, AND GUESS DEFINITIONS **301**

A.1. DEFINITIONS OF REMEMBER, KNOW, FAMILIAR, AND GUESS FROM EXPERIMENTS 2.2 AND 2.3.301

A.2. DEFINITIONS OF REMEMBER, KNOW, FAMILIAR, AND GUESS FROM EXPERIMENTS 3.1 AND 3.2.302

A.3. DEFINITIONS OF REMEMBER, KNOW, FAMILIAR, AND GUESS FROM EXPERIMENTS 3.3, 4.1, AND 4.2 303

APPENDIX B: STIMULI FROM ALL EXPERIMENTS **304**

B.1. CUE WORDS AND JUSTIFICATION STATEMENTS USED AS STIMULI IN EXPERIMENTS 2.1 TO 2.4. .. 304

B.2. FORENAME-SURNAME PAIRINGS USED IN EXPERIMENTS 3.1 AND 3.2. 324

B.3. WORD LISTS USED IN EXPERIMENT 3.3. 325

B.4. WORD LISTS USED IN EXPERIMENTS 4.1, 4.2, AND 4.3. 329

APPENDIX C: FORMULAE FOR YES/NO AND 2AFC D-PRIME CALCULATIONS **330**

APPENDIX D: FORMULAE FOR COMPARING CORRELATIONS **331**

APPENDIX E: FORMULA FOR GAMMA CORRELATION **332**

APPENDIX F: DECISION CRITERIA QUESTIONNAIRE FOR EXPERIMENT 2.4 .... **333**

APPENDIX G: ITEMS ANALYSIS, EXPERIMENT 2.4 .... **337**

APPENDIX H: ALTERNATIVES TO FIGURES 3.8 AND 3.9, EXPERIMENT 3.3 .... **338**

APPENDIX I: ALTERNATIVE TO FIGURE 4.15, EXPERIMENT 4.2 .... **340**

LIST OF FIGURES

Figure 1.1. Equal-Variance Signal Detection (EVSD) model (left panel) and Unequal-Variance Signal Detection (UVSD) model (right panel) distribution curves for New and Old items. 25

Figure 1.2. Example ROCs in probability space (left panel) and z-space (right panel) after Parks and Yonelinas (2008). 36

Figure 2.1. Proportion of Gardiner et al.'s (1998) Remember, Know, and Guess justification statements assigned to each level of confidence. 72

Figure 2.2. Proportion of original Know responses that were assigned to Know and Familiar at each confidence level. 73

Figure 2.3. Proportion of justifications of the four original subjective experience categories that were categorised as Remember, Know, Familiar, and Guess. 80

Figure 2.4. Proportion of REMEMBER justifications given High, Medium, and Low ratings of confidence in Experiment 2.1 that were assigned to Remember, Know, Familiar, and Guess. 81

Figure 2.5. Proportion of KNOW justifications given High, Medium, and Low ratings of confidence in Experiment 2.1 that were assigned to Remember, Know, Familiar, and Guess. 83

Figure 2.6. Proportion of FAMILIAR justifications given High, Medium, and Low ratings of confidence in Experiment 2.1 that were assigned to Remember, Know, Familiar, and Guess. 85

Figure 2.7. Proportion of GUESS justifications given High, Medium, and Low ratings of confidence in Experiment 2.1 that were assigned to Remember, Know, Familiar, and Guess. 86

Figure 2.8. Proportion of REMEMBER justifications matched with Very High, High, Medium, and Low confidence values that were assigned to Remember, Know, Familiar, and Guess. 96

Figure 2.9. Proportion of KNOW justifications matched with Very High, High, Medium, and Low confidence values that were assigned to Remember, Know, Familiar, and Guess. 96

Figure 2.10. Proportion of FAMILIAR justifications matched with Very High, High, Medium, and Low confidence values that were assigned to Remember, Know, Familiar, and Guess.. 97
Figure 2.11. Proportion of GUESS justifications matched with Very High, High, Medium, and Low confidence values that were assigned to Remember, Know, Familiar, and Guess.. 97
Figure 2.12. Proportion of justifications at each level of manipulated confidence that were categorised as Remember, Know, Familiar, and Guess irrespective of original subjective experience category. 98
Figure 2.13. Proportion of justifications of the four original subjective experience categories that were categorised as Remember, Know, Familiar, and Guess regardless of manipulated confidence level.. 99
Figure 3.1. Proportion of correct responses assigned to Remember, Know, Familiar, and Guess by participants in the Pre-Exposure and No Pre-Exposure conditions. 138
Figure 3.2. Proportion of forenames assigned to Remember, Know, Familiar, and Guess in the two conditions that had previously been correctly recognised. 140
Figure 3.3. Mean correct recognition RTs by later subjective experience category. 142
Figure 3.4. Proportion of correct responses assigned to Remember, Know, Familiar, and Guess by participants in the Pre-Exposure and No Pre-Exposure conditions. 148
Figure 3.5. Proportion of forenames assigned to Remember, Know, Familiar, and Guess in the two conditions that had previously been correctly recognised. 149
Figure 3.6. Proportion of correct responses assigned to Remember (R), Know (K), Familiar (F), and Guess (G) by participants in the Pre-Exposure and No Pre-Exposure conditions of the Yes/No task and the 2AFC task. 161
Figure 3.7. Proportion of correct responses assigned to Remember (R), Know (K), Familiar (F), and Guess (G) by participants in the Pre-Exposure and No Pre-Exposure conditions of the Yes/No task split by which order they performed the conditions in. 163
Figure 3.8. Proportion of items assigned to Remember (R), Know (K), Familiar (F), and Guess (G) in the in the Pre-Exposure and No Pre-Exposure conditions of the Yes/No task and the 2AFC task that had previously been correctly recognised. 165

Figure 3.9. Mean correct recognition RTs by later subjective experience category, Remember (R), Know (K), Familiar (F), and Guess (G), in the in the Pre-Exposure and No Pre-Exposure conditions of the Yes/No task and the 2AFC task........ 168
Figure 4.1. Mean proportion of correct responses assigned at First and Second Judgments to subjective experience categories Remember (R), Know (K), Familiar (F), and Guess (G), and confidence levels High, Medium, Low, and None. 195
Figure 4.2. Mean proportion of items assigned to Remember (R), Know (K), Familiar (F), and Guess (G) and High, Medium, Low, and None levels of confidence that had previously been correctly recognised........ 197
Figure 4.3. Mean proportion of items assigned to each level of confidence at First Judgment that were then assigned to each subjective experience category at Second Judgment........ 199
Figure 4.4. Mean proportion of items assigned to each level of confidence at Second Judgment that had previously been assigned to each subjective experience category at First Judgment. 200
Figure 4.5. Mean proportion of items assigned to each subjective experience category at First Judgment that were then assigned to each level of confidence at Second Judgment. 202
Figure 4.6. Mean proportion of items assigned to each subjective experience category at Second Judgment that had previously been assigned to each level of confidence at First Judgment. 203
Figure 4.7. Mean confidence ratings given to items assigned to Remember, Know, Familiar, or Guess. 205
Figure 4.8. Mean correct recognition RTs split by later response category. 206
Figure 4.9. Mean judgment RTs (ms) by judgment type and judgment order. 208
Figure 4.10. Mean judgment RTs (ms) by response category. 209
Figure 4.11. Mean proportion of correct responses assigned at First and Second Judgments to subjective experience categories Remember (R), Know (K), Familiar (F), and Guess (G) in Experiment 4.2 and confidence levels High, Medium, Low, and None in Experiment 4.3. 221
Figure 4.12. Mean proportion of items assigned to Remember (R), Know (K), Familiar (F), and Guess (G) in Experiment 4.2 and High, Medium, Low, and None levels

of confidence in Experiment 4.3 that had previously been correctly recognised. 223
Figure 4.13. The influence of source and subjective experience judgments on one another, Experiment 4.2. 224
Figure 4.14. Source accuracy split by subjective experience judgment, Experiment 4.2. 226
Figure 4.15. The influence of source and confidence judgments on one another, Experiment 4.3. 228
Figure 4.16. Source accuracy split by confidence judgment, Experiment 4.3. 229
Figure 4.17. Mean correct recognition RTs split by later response category in Experiment 4.2. 231
Figure 4.18. Mean correct recognition RTs split by later response category in Experiment 4.3. 232
Figure 4.19. Mean source judgment RTs (ms) by accuracy and judgment order in Experiments 4.2 and 4.3. 233
Figure 4.20. Mean judgment RTs (ms) by judgment type and judgment order. 234
Figure 4.21. Mean judgment RTs (ms) by response category. 235
Figure 4.22. Mean time to make correct source judgment split by the subjective experience category (Experiment 4.2, left panel) or confidence level (Experiment 4.3, right panel) the item was assigned to. 237

List of Tables

Table 2.1. Example Remember justification statements for each classification type, from Gardiner et al. (1998). 64
Table 2.2. Example Know and Guess justification statements from Gardiner et al. (1998). 65
Table 2.3. Mean confidence ratings assigned to Gardiner et al.'s (1998) Remember, Know, and Guess justifications (and standard deviations). 71
Table 2.4. Mean confidence ratings for Remember, Know, Familiar, and Guess justifications (and standard deviations). 72
Table 2.5. Maximum, minimum, and mean confidence ratings for the selected justifications from each subjective experience category. 77
Table 2.6. Means and ranges of confidence values for the justifications from each subjective experience category. 93
Table 2.7. Example data of Know and Familiar justifications sorted into Type A and Type B. 109
Table 2.8. Mean gamma correlation coefficients (and standard deviations) calculated using different methods of mapping Type A and B to Know and Familiar definitions. 110
Table 2.9. Mean scores and standard deviations for ratings of ease, similarity, and other peoples' sorting. 111
Table 2.10. Decision criteria for statement sorting and the percentage of participants who endorsed each option. 112
Table 2.11. Mean information ratings (and standard deviations) split by whether the participant had endorsed 'Recalled more information' as a decision criterion. 113
Table 3.1. Correlations between memory accuracy and proportion of items assigned to subjective experience categories in each condition. 141
Table 3.2. Correlations between memory accuracy and proportion of items assigned to subjective experience categories in each condition. 151
Table 3.3. Means (and standard deviations) of recognition performance and *d'* values in each of the blocks. 159

Table 3.4. Means (and standard deviations) of recognition performance in Pre-Exposure and No Pre-Exposure conditions split by which exposure condition was performed first or second. 160
Table 3.5. Correlations between memory accuracy and proportion of items assigned to subjective experience categories in each condition. 166
Table 4.1. Mean memory measures (and standard deviations) in Experiments 4.2 and 4.3. 220
Table 5.1. Main findings from each experiment. 249
Table 5.2. Main findings for the separated categories of Know and Familiar. 263

ABBREVIATIONS

AFC	Alternate forced choice
EEG	Electroencephalography
ERP(s)	Event-related potential(s)
EVSD	Equal-variance signal detection
F	Familiar
FA(s)	False alarm(s)
fMRI	Functional magnetic resonance imaging
G	Guess
K	Know
LOP	Levels of processing
ms	Milliseconds
MTL	Medial temporal lobe
PDP	Process-dissociation procedure
R	Remember
ROC	Receiver-operating characteristic
RT(s)	Reaction time(s)
SAC	Source activation confusion
SD	Standard deviation
SDT	Signal detection theory
SeM	Standard error of the mean
SPT	Subject performed task
UVSD	Unequal-variance signal detection

1. General Introduction

"I enter a friend's room and see on the wall a painting. At first I have the strange, wondering consciousness, 'surely I have seen that before', but when or how does not become clear. There only clings to the picture a sort of penumbra of familiarity, - when suddenly I exclaim: 'I have it, it is a copy of part of one of the Fra Angelicos in the Florentine Academy - I recollect it there!'"

(William James, 1890, p. 658)

1.1. Overview

The above quote illustrates how people can have conscious awareness of their memory processes. Recognition can either occur with recollection, where the object of recognition evokes retrieval of other related thoughts, feelings or memories; or recognition can occur without triggering any such associations and instead is accompanied by feelings of familiarity (Gardiner, 2000). Importantly for the study of recognition memory, people are also able to consciously appreciate the differences between these processes (Wixted & Mickes, 2010).

Psychology has sought to explain these two forms of recognition through a number of different models and approaches including the Remember-Know paradigm (Tulving, 1985) which is used in all experiments in this thesis. This chapter will introduce recollection and familiarity and the Remember-Know paradigm before describing important gaps in the existing literature and presenting a rationale for the programme of research presented in this thesis. In brief, this thesis aimed to further elucidate and classify the subjective experiences contributing to recognition memory, with a view to examine how subjective experience can be sensitive to and manipulated by experimental factors. This is an issue of value to those interested in how subjective experience might be used to understand recognition memory failure and strategy use in memory retrieval (see Section 1.3 on applications).

Human memory is a dynamic system relying on separate stores of information and different strategic and automatic retrieval processes. At a basic level, it is possible to differentiate between episodic (experience-based) and semantic (fact-based) memory stores and longer term and shorter term stores (Atkinson & Shiffrin, 1968; Tulving, 1972). Within the focus of this thesis, episodic memory, it is possible that previous experience is retrieved either through recall or recognition. Recall is a relatively effortful process whereby the instances of a previous experience are reproduced in response to cues, whereas in contrast, recognition (the focus of the current thesis) is a process whereby a previously encountered stimulus is re-experienced and is endorsed as either being something previously experienced (Old) or rejected as being novel (New). Within recognition memory research, a significant development has been the use of subjective experience to classify different forms of recognition memory.

1.2. Recollection and familiarity; Remembering and Knowing

As illustrated in the quote at the start of this chapter, recognition can be accompanied by different subjective experiences. Theoretical accounts detailing different memory processes, types of memory, and states of awareness that have been suggested to underlie different recognition experiences are detailed in Section 1.4; however, to foreshadow these theoretical accounts, the central idea is that recognition is sometimes accompanied by *recollective experience*. Recollective experience involves retrieval of details of a past encounter accompanied by a sensation of mental time travel or a feeling of oneself in the past (Gardiner, 2008). For example, when visiting a city for the second time you might chance upon a shady park and recall that last time you visited that city you came to the same park but it was a miserable rainy day and you could not sit and enjoy the surroundings. If recognition is not accompanied by recollective experience the subjective feeling may instead be one of *familiarity*; the shady park may feel familiar but no additional contextual details are retrieved about the previous visit to the park. Or you may *know* that you have been to the park before but it does not feel familiar and you do not have any recollective experience for the surroundings.

Theoretical accounts have conceptualised subjective experiences as relating to the processes of *recollection* and *familiarity* (e.g., Mandler, 1980; Jacoby, 1991; Yonelinas, 1997) or to the states of awareness of *remembering* or *knowing* (e.g., Tulving, 1985; see

Section 1.4 for full details of these different accounts). This latter conceptualisation included the development of the Remember-Know paradigm where participants are instructed to categorise recognised items as either Remembered, where recognition is accompanied by recollective experience, or Known, where recognition is not accompanied by recollective experience and instead is accompanied by a feeling of familiarity (Yonelinas, 2002). The differentiation of *knowing* and *familiarity* as independent subjective experiences is one of the key themes of this thesis and, while the majority of previous research reviewed in this chapter has conflated these two experience types, the importance of their separation is discussed in Section 1.6.5.3. The Remember-Know paradigm is the central focus of this thesis and detailed discussion of the dissociation of Remember and Know responses is undertaken in Section 1.6. However, the importance of subjective experience for understanding memory processes and behaviours is first reviewed.

1.3. Implications and applications

When introspective techniques were banished in favour of more objective experimental procedures, psychology concentrated on examining memory performance through behavioural measures (cf. James, 1890, and Ebbinghaus, 1885). However, the re-emergence of subjective experience as a valid means for exploring cognitive processes since the 1960s has enabled memory researchers to examine the feelings, thoughts, and cognitive control processes related to behavioural measures of memory (Gardiner, 2008). Two central questions of interest in this area are: What do subjective feelings tell us about our memory processes? And how do these feelings influence our subsequent behaviours?

In recent decades, subjective experience has been explored in nearly every area of memory research. In autobiographical memory, retrieval accompanied by recollective experience has been found to differentiate between true and false memories (Conway, Collins, Gathercole, & Anderson, 1996) as well as whether childhood memories are viewed from an observer or field perspective (Crawley & French, 2005). Impairments in recollective experience have also been linked to the lack of autobiographical memory specificity which characterise disorders such as depression (Ramponi, Barnard, & Nimmo-Smith, 2004). Thus, recollective experience is important to many aspects of autobiographical memory and, as such, recollective experience has been identified as one of the defining characteristics of episodic memory (Conway, 2005; Conway & Williams, 2008).

Recollective experience has also been demonstrated to be of value in applied settings. In eyewitness memory, retrieval accompanied by recollective experience has been found to distinguish correct from incorrect line-up identifications (Palmer, Brewer, McKinnon, & Weber, 2010); and has been demonstrated to be the element of recognition crucially sensitive to the binding of faces to contextual information in the recognition of people (Gruppuso, Lindsay, & Masson, 2007). Differences in recollective experience are also of interest in terms of how people learn information in academic settings. Mirandola, Del Prete, Ghetti, and Cornoldi (2011) explored recollective experience for sentences from a text passage in adolescents with learning difficulties and found that poor learners reported lower levels of recollective experience for sentences than did participants without learning difficulties. Furthermore, as detailed in Section 1.6.5.3 below, Conway, Gardiner, Perfect, Anderson, and Cohen (1997) and Herbert and Burt (2001, 2003, 2004) have demonstrated that students' retrieval shifts from Remember to Know, or from episodic retrieval to conceptual organisation of knowledge across time and this shift is associated with improvements in academic performance.

Reduced levels of recollection have also been implicated as one of the fundamental changes associated with decreased episodic memory performance in normal aging (Friedman & Trott, 2000; Java, 1996; Norman & Schacter, 1997; Parkin & Walter, 1992; Perfect & Dasgupta, 1997; Perfect, Williams, & Anderton-Brown, 1995; Souchay, Moulin, Clarys, Taconnat, & Isingrini, 2007); and research has started to explore whether recollection can be improved in older adults (Jennings & Jacoby, 2003). Remembering has also been shown to be much reduced in conditions such as Alzheimer's disease, amnesia, schizophrenia, and autism spectrum disorders, while levels of Knowing remain unaffected or are affected to a much lesser extent than Remembering (Bowler, Gardiner, & Grice, 2000; Dalla Barba, 1993, 1997; Huron et al., 1995; Knowlton & Squire, 1995; Schacter, Verfaellie, & Anes, 1997; Tanweer, Rathbone, & Souchay, 2009). Together these examples demonstrate how renewal of interest in subjective experience has led to increases in what we know about memory behaviours and processes in a wide variety of domains. Exploration of subjective experience therefore has both widespread implications for theory development and applications for improving memory.

1.4. THEORETICAL ISSUES

1.4.1. Dual-process versus single-process accounts of recognition memory

Dual-process models assume that two distinct processes or forms of memory underlie recognition and that successful recognition is determined by the relative contributions of both processes. Conversely, single-process accounts propose that recognition relies on only one continuous dimension of familiarity or memory strength, and successful recognition is determined by the strength of this single dimension. The following sections provide brief overviews of single- and dual-process accounts and highlight the major points of agreement, disagreement, and debate between the various models. For comprehensive reviews of the literature which summarize the opposing viewpoints see Yonelinas (2002), Diana, Reder, Arndt, and Park (2006), Dunn (2004, 2008), Parks and Yonelinas (2007), and Wixted and Stretch (2004).

1.4.2. Dual-process models

1.4.2.1. Mandler

George Mandler and colleagues (Mandler, 1979, 1980, 1991, 2008; Mandler & Boeck, 1974; Mandler, Pearlstone, & Koopmans, 1969) argue that memory consists of a fast perceptual matching process and a slower retrieval process. In the process of perceptual matching, if an item matches representations in memory this activation gives rise to subjective feelings of familiarity. Contrastingly, the slower retrieval process is recall-like and is based on an active search of memory. According to this model, familiarity will be higher if an item has been repeatedly and/or recently experienced as the item will have been more strongly integrated in memory. This familiarity process is considered to support recognition memory judgments and performance on implicit memory tasks such as word stem completion (Mandler, 1991). Contrastingly, a recollection-based search process is assumed to act if experience with an item has led to the item being elaborated into semantic networks in memory. These organisational structures can be searched and stored information may be accessed from a variety of directions. This recollection-based search process is assumed to support both recognition and recall performance. Though Mandler et al. (1969) initially suggested that active retrieval search processes are only initiated if the result of the familiarity process is ambiguous, later conceptualisations of the model identified familiarity and recollection as independent processes which act in parallel, but with familiarity typically being faster than recollection and therefore 'winning the race' in recognition memory tests (Mandler, 1980, 2008). The issue of the time course of subjective experiences is of critical importance to experiments presented in Chapters 3 and 4 of this thesis and research related to the speed of recognition decisions is discussed there (e.g., Sections 3.2.3.4 and 4.1.3).

1.4.2.2. Jacoby

Jacoby's model (Jacoby, 1983, 1984, 1991; Jacoby & Dallas, 1981; Jacoby & Kelley, 1992; Jacoby, Kelley, & Dywan, 1989; Jacoby & Witherspoon, 1982; Kelley & Jacoby, 1990; Whittlesea, Jacoby, & Girard, 1990) also focuses of speed of processing as a method of differentiating between recognition processes. This model stipulates that recognition of an item can be based on an assessment of processing fluency, seen as a relatively automatic process; or recollection of contextual or elaborative information linked to the item, conceived as an analytic, consciously controlled process. In this model, feelings of familiarity arise when fluent processing of an item is attributed to past experience with the item. Fluency of processing is also not limited to perceptual fluency; feelings of familiarity can also arise from increases in conceptual fluency (e.g., enhanced processing of the meaning of an item; Jacoby, 1984; 1991; Jacoby & Kelley, 1992). Similar to Mandler's (1980) later modifications of his model, Jacoby conceptualised familiarity and recollection as independent processes that operate in parallel but with familiarity typically being a faster process than recollection (Jacoby, 1991). Jacoby went on to develop the process-dissociation procedure (PDP) as a methodology for estimating the relative contributions of recollection and familiarity to recognition (Jacoby, 1991); this paradigm is discussed further in Section 1.5.2.1.

1.4.2.3. Tulving

Endel Tulving and colleagues (Nyberg, Cabeza, & Tulving, 1996; Tulving, 1982, 1985; Tulving & Markowitsch, 1998; Tulving & Schacter, 1990; Wheeler, Stuss, & Tulving, 1997) suggested there are dual *states of awareness* underlying memory performance rather than dual *processes*. This conceptualisation developed from Tulving's distinction between episodic and semantic memory systems (Tulving, 1972). The episodic system is stipulated to contain records of personally experienced events which, when retrieved, give rise to the state of autonoetic consciousness and the subjective experience of 'remembering'.

Conversely, the semantic system stores general knowledge about the world, and when information is retrieved from this system it comes to mind with noetic consciousness, and the accompanying subjective feeling is one of 'knowing' or a feeling of familiarity without recollection (Yonelinas, 2002). The episodic memory system is considered to underlie recall, while both episodic and semantic systems are utilised for recognition. During memory retrieval, the two systems are assumed to be independent but operating in parallel as information can be retrieved from either system individually (Tulving, 1985, 1995; Tulving & Markowitsch, 1998).

Tulving (1985) was the first of his contemporaries to explore memory retrieval through subjective report. He asked participants to report their state of awareness for each retrieved item by making a 'Remember' or 'Know' judgment. Items were categorised as Remembered when the participant retrieved some memory of something they had thought or experienced at the time of encoding and categorised as Know when the participant was aware that the item had been on the study list but could not recall anything experienced for the item at that time (Gardiner & Richardson-Klavehn, 2000; Tulving, 1985). While some view the subjective states of Remembering and Knowing as orthogonal to the underlying processes of recollection and familiarity (for example see Gardiner, 2000, and Wixted & Mickes, 2010), the majority of studies utilising the Remember-Know paradigm assume that the two categories map directly onto the underlying processes (Yonelinas, 2002). Furthermore, experiments using the Remember-Know procedure have demonstrated similar patterns of findings to other methodologies exploring the contributions of recollection and familiarity processes to recognition (e.g., Yonelinas, 2001a). The Remember-Know paradigm forms the basis for all the experiments presented in this thesis and therefore is discussed in more detail in Section 1.6.

1.4.2.4. Yonelinas

The dual-process signal detection model (DPSD) put forward by Yonelinas and colleagues argues that recollection and familiarity are qualitatively different processes that each provide different information and levels of confidence to recognition (Dobbins, Kroll, Yonelinas, & Liu, 1998; Yonelinas, 1994, 1997, 1999, 2001a, 2001b, 2002; Yonelinas, Aly, Wang, & Koen, 2010; Yonelinas, Dobbins, Szymanski, Dhaliwal, & King, 1996; Yonelinas, Kroll, Dobbins, Lazzara, & Knight, 1998; Yonelinas, Kroll, Dobbins, & Soltani, 1999). The DPSD conceptualisation of familiarity builds on single-process signal-detection models and receiver-operating characteristic (ROC) curve methodologies (see below, Sections 1.4.3 and 1.5.2.2) whereby familiarity is considered to reflect a quantitative measure of memory strength that is temporarily increased when an item is studied. In contrast, recollection is conceptualised as a threshold process where qualitative information about a previous event is either retrieved or not. If qualitative information about the study episode is not retrieved, and therefore the threshold for recollection is not met, recognition must rely only on familiarity. In line with the previous models of Jacoby and Mandler, the DPSD model also assumes that recollection and familiarity are independent and are initiated in parallel, and familiarity is expected to be faster than recollection (Yonelinas & Jacoby, 1994, 1996; Yonelinas, 2002).

In relation to recognition confidence, the DPSD assumes that familiarity can support a wide range of confidence responses, while recollection is expected to support relatively high confidence recognition decisions only. This is reflected in the model's assumption that all items will evoke a familiarity signal, but only some items will be recollected. Here recollection is conceptualised as a threshold process, although the model does assume that participants can recollect different types or amounts of information experienced at encoding. Whilst this may not seem to reflect a threshold process, it is conceptualised as such because of the fact that recollection can sometimes fail and no contextual information will be retrieved. A further assumption of the DPSD model is that recollection is assumed to have a distribution of strength, but the model does not specify what kind of distribution it is and therefore recollection is measured as the probability of success (Parks & Yonelinas, 2008). The DPSD was the first model to suggest a hybrid of SDT measures for familiarity and a threshold process for recollection, though other hybrid models have started to be put forward in the last decade, some of which conceptualise recollection as a continuous process and provide estimates for its distribution (Parks & Yonelinas, 2008; Wixted & Mickes, 2010; see Section 1.4.5).

1.4.3. Single-process models

Single-process accounts of recognition memory are based on signal detection theory (SDT). Two core assumptions of single-process accounts are that items in a memory task can be ordered along a single continuum of familiarity or strength of evidence and that the strength distribution curves for old and new items are overlapping and Gaussian (or normal) in shape, as illustrated in Figure 1.1 (Dunn, 2004; Hirshman & Master, 1997; Parks & Yonelinas, 2008). Researchers who use SDT modelling tend to retrospectively fit models to data sets to examine which model parameters accurately describe the patterns of data obtained; further predictions are then able to be made from the best-fit model (e.g., Dunn, 2004, 2008; Rotello, Macmillan, & Reeder, 2004).

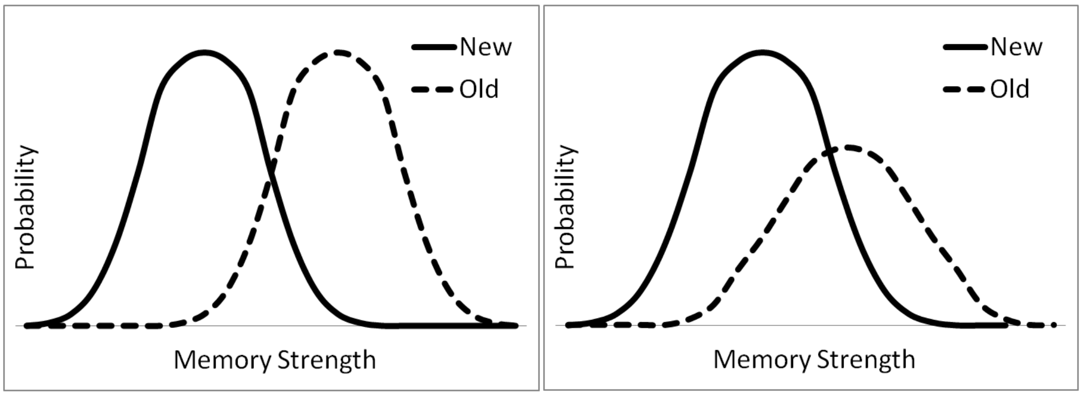

Figure 1.1. Equal-Variance Signal Detection (EVSD) model (left panel) and Unequal-Variance Signal Detection (UVSD) model (right panel) distribution curves for New and Old items.

Within SDT, where a person places their recognition decision criterion along the memory strength axis determines the probability of their correctly recognising old items while not making false alarms for new items, and accuracy is measured by *d′* which is the difference in average strength of the new and old item distributions[1].

A third assumption of single-process accounts comes into play when SDT is used to interpret Remember-Know responses. This assumption is that the proportions of Remember and Know responses are determined by the placement of two criteria on the familiarity or strength-of-evidence axis instead of one. A criterion placed low on the recognition decision axis is used to distinguish old (including both Remember and Know) from new responses, whilst a criterion placed higher on the decision axis is used to distinguish Remember from Know responses (Donaldson, 1996; Dunn, 2004; Hirshman & Master, 1997; Inoue & Bellezza, 1998; Parks & Yonelinas, 2008).

[1] *d′* values are reported for all recognition memory experiments in this thesis and the full formula for calculating *d′* is shown in Appendix C.

As is shown in Figure 1.1, two variants of SDT are the Equal-Variance Signal Detection (EVSD) model and Unequal-Variance Signal Detection (UVSD) model. The EVSD model is the simplest SDT model and it assumes that the variance of old and new items' strength distributions are equal. In contrast, the UVSD model assumes that old and new items' strength distributions differ in variance. This is suggested to occur because at study some items are likely to increase in strength more than others meaning that distribution for studied items will be more variable than the distribution for new items (Parks & Yonelinas, 2008). In this model two parameters are measured: the distance between the means of the distributions (equivalent to *d'*) and the difference between the variances of the distributions.

In their review of this literature, Parks and Yonelinas (2008; and Yonelinas & Parks, 2007) evaluate the underlying assumptions of these two signal detection models as well as other more recent variants including threshold and mixed models such as the hybrid Dual-Process Signal Detection model discussed in Section 1.4.2.4. They also discuss methodological issues involved in analysing results from ROC curves, the shapes of which are important for determining which models can account for which data. Their discussion of ROC curves is returned to in Section 1.5.2.2 below. A comprehensive review of these models and their mathematical differences is beyond the scope of the current thesis. Furthermore, the thesis did not aim to discriminate between single- and dual-process accounts, principally because the interest is in how different forms of subjective experience in recognition differ. Thus, a theory-neutral position is taken whereby confidence (as a report of trace strength) and recollective experience (as a qualitative state) are compared within individuals. However, the following sections consider some key points regarding the debate between dual- and single-process accounts relating to subjective experiences and their relationship with recognition, since it is likely that studies comparing confidence and recollection may make a modest contribution to this debate.

1.4.4. Single-process accounts of Remember-Know data

A critical issue in recognition memory is whether the subjective states used to differentiate dual-process accounts of recognition have a sound basis. To this end, a number of meta-analyses have examined recognition memory performance to examine the differences between memory strength for the different subjective experience responses. If SDT models can accommodate dual-process data, it suggests that a more parsimonious account of human recognition is to consider a single, strength-based process.

Donaldson (1996) initiated this SDT versus Remember-Know debate when he performed a meta-analysis of 80 experimental conditions to examine whether the recognition accuracy of the Remember criterion was significantly different to the accuracy of the recognition (Old-New) criterion. He did not find a significant difference. This supports a signal detection interpretation of Remember-Know judgments as it suggests that whether Know responses are included under a lenient decision criterion or whether only Remember responses are included using a strict criterion this does not result in differences in recognition sensitivity and therefore all recognition decisions must be based on the same underlying memory trace (Gardiner, Ramponi, & Richardson-Klavehn, 2002). However, Gardiner and Gregg (1997) determined that Donaldson's values of *A′* (.86 for recognition and .83 for Remember) were significantly different when compared using a sign test, and thus began a pattern of conflicting results from meta-analyses.

Gardiner et al. (2002) performed a similar meta-analysis to that of Donaldson (1996) though data from none of the included 86 empirical conditions overlapped with Donaldson's as Gardiner et al. included only experiments that had allowed a Guess response (see Section 1.6.5.2) while Donaldson had excluded these. This addition was crucial as it meant that three levels of response criterion could be compared – a Guess criterion (which includes all recognised items), a Know criterion and a Remember criterion. Contrary to the analysis conducted by Donaldson, Gardiner et al. found that the strength of the memory trace was not equal at these different decision criteria points; *A′* was .799 for recognition (Remember + Know + Guess), .813 for Remember plus Know responses, and .787 for Remember responses alone. Gardiner et al. (2002) concluded that "the addition of know responses to remember responses increased memory strength and therefore that these responses reflect an additional source of memory, not merely more lenient response criteria" (p. 91; and see results from Gardiner & Gregg, 1997, and Gardiner & Conway, 1999). However, the underlying assumptions, model parameters, and statistical methods employed in dual-process interpretations of signal-detection analyses such as that of Gardiner et al. were subsequently questioned by Dunn (2004, 2008). In his 2004 paper, Dunn detailed five statistical arguments suggested by dual-process protagonists as to why SDT could not be used to interpret Remember-Know data. Dunn reasoned that these five arguments were flawed and concluded that there was no evidence from Remember-Know paradigms that was inconsistent with an SDT interpretation. He later furthered this argument by involving state-trace analysis in a meta-analysis in an aim to determine the number and organisation of mediating psychological variables involved in recognition memory. He again concluded that a unidimensional model fit the data better than a bidimensional model (Dunn, 2008).

In addition to the basic underlying premises of one or two processes and the complex statistical evidence put forward by the opposing viewpoints to support their premise, a further conceptual difference between the view held by single- and dual-process protagonists Dunn and Gardiner is whether recollection and familiarity are antecedent to or resultant from recognition decisions. Dunn (2008) suggests that even if it is agreed that Remember and Know responses are associated with different kinds of subjective experience, it does not automatically follow that these experiences are what give rise to the type of judgment made; instead they could be the result of what response is made. Dunn goes on to provide the following example “following evaluation that a test item has sufficient strength of evidence to be classified as old, an R response may also be interpreted, on the basis of the level of this strength of evidence, as an inference that some details of the earlier study episode could be retrieved if required” (2008, p. 442). By this way of thinking, Remember and Know are inferences based on the strength of evidence of an item and can be viewed as meta-memory judgments relating to the type of accompanying information that is expected to be retrieved. In contrast, as proposed by Tulving (1985) and Gardiner “…it is surely the subjective state of awareness that gives rise to confidence in memory, not confidence that gives rise to the state of awareness” (Gardiner, 2001, p. 1356). This is a more fundamental conceptual difference of opinion than arguments based around statistical sensitivity and one of more importance if your concern is with understanding the states of awareness and conscious experiences associated with Remember and Know responses, as this thesis is.

1.4.5. Recent additions to the theoretical debate

In the last decade a flurry of other dual-, single-, mixed and hybrid process models have been put forward. New dual-process models include the source of activation confusion (SAC) model proposed by Reder and colleagues (Diana, Reder, Arndt, & Park, 2006; Reder et al., 2000), the theory of distributed associative memory (TODAM)–based model proposed by Murdock (2006), sum-difference theory of Remembering and Knowing (STREAK) proposed by Rotello et al. (2004), and models proposed by Wixted and Stretch (2004) and Hintzman (2001). As discussed by Macmillan and Rotello (2006; Rotello & Macmillan, 2006), these models can be classified according to which of three different decision rules are used. The SAC (Diana et al., 2006; Reder et al., 2000) and TODAM (Murdock, 2006) models use the “process-pure” rule whereby the two processes underlying recognition decisions are conceptualised as being uniquely relevant to either Remember or Know responses and operating in hierarchical fashion. If there is sufficient information available from the first process (recollection) to determine recognition, then a Remember response is made; otherwise, if there is sufficient information available from the second process (familiarity), then a Know response is made; otherwise, a new response is made (Dunn, 2008; Macmillan & Rotello, 2006; Rotello & Macmillan, 2006).

The second decision rule is the “sum-difference” rule used in Rotello et al.’s (2004) STREAK model. According to this rule, Old-New decisions are based on global memory strength, which is related to the *sum* of evidence from the processes of recollection and familiarity. Contrastingly, Remember-Know decisions are based on the *difference* between the specific memory strength available from recollection, which should be relatively great compared with global memory strength, and the specific memory strength available from familiarity, which should be relatively weak. Thus in this model, the weighted sum of global and specific strengths results in a recognition decision, and the weighted difference of strengths results in either a Remember response or a Know response.

The third type of decision rule characterizes the models proposed by Wixted and Stretch (2004) and Hintzman (2001). In these models, it is suggested that people combine information from different sources into a single strength-of-evidence value on which both recognition and Remember-Know decisions are based. This results in models which are mathematically equivalent to the single-process model of the Remember-Know task first proposed by Donaldson (1996). Although the models of Wixted and Stretch (2004) and Hintzman (2001) agree with the dual-process assumption that recognition memory judgments are based on two different processes or sources of information, they reject the other assumption that recognition decisions can be made on the basis of either recollection *or* familiarity. Instead, they suggest that the two sources of information are pooled and mapped onto a single decision axis. A further suggestion from these models and others is that recollection is a graded process existing on a continuum like familiarity, as opposed to a threshold or all-or-none process as was originally suggested by the DPSD model (e.g., Mickes, Johnson, & Wixted, 2010; Mickes, Wais, & Wixted, 2009; Onyper, Zhang, & Howard, 2010; Parks & Yonelinas, 2007; Rotello et al., 2004; Slotnick, 2010; Slotnick & Dodson, 2005; Wixted, 2007, 2010; Wixted & Mickes, 2010; Wixted & Stretch, 2004).

Whilst these recent models can be differentiated in terms of the decision rules they employ, most of them provide a reasonably good fit to the data (Gardiner, 2008). Because of this, and the fact that the specifics of the models involve such complexity and technical sophistication, Gardiner (2008) highlighted that it has become increasingly difficult to see how to distinguish between them empirically. While these recent additions to the debate are interesting, and will be revisited in the General Discussion, the current thesis is not concerned with providing evidence which supports one or other of the models. Rather the experiments presented here take Remember-Know judgments at face value and are interested in understanding how people make and understand judgments of subjective experience. This move 'back' towards understanding Remembering and Knowing at a more conceptual level was advocated by both Dunn (2004) and Gardiner (2008).

1.5. MEASURING RECOLLECTION AND FAMILIARITY

To support the idea that multiple processes are involved in recognition memory, various measures have been put forward as being able to dissociate the processes of recollection and familiarity. These fall into two main camps: Task-dissociation methods and process-dissociation methods.

1.5.1. Task-dissociation methods

Task-dissociation methods aim to find a task or experimental condition for which the results dissociate the processes of recollection and familiarity. To find this dissociation the task or condition has to produce results that are different from a standard recognition task where both recollection and familiarity processes are involved in performance. Inferences can then be made about the relative contributions of recollection and familiarity to the task or condition (Yonelinas, 2002). Task-dissociation methods include response speed methods and associative versus single-item recognition.

1.5.1.1. Response-speed methods

Many dual-process models, including those of Mandler, Jacoby, and Yonelinas (see Section 1.4.2) consider familiarity to be a rapid and automatic process with relatively low demands on cognitive resources, while recollection is a slower and more effortful. Because of this, response speed has been suggested as a way to dissociate the two processes (Yonelinas, 2002). Response speed can be measured by simply separately analysing recognition performance for fast and slow recognition responses or by setting a response deadline and comparing performance in this speeded condition to performance in a non-speeded condition. In both these methods, fast responses are thought to reflect familiarity, while slower responses (when allowed) are thought to involve more contribution from recollection (Yonelinas, 2002).

Yonelinas (2002) reviewed the literature that has used response-speed methods and concluded that familiarity is indeed available earlier and contributes to performance earlier in the time course than recollection does. For example, in studies where participants are instructed to exclude lures that are related in some way to actually studied items (either visually, semantically, or previously paired with), accuracy in rejecting these lures has been shown to improve as response deadline is increased. At short deadlines, participants have a tendency to accept the lures due to familiarity but given additional retrieval time they are able to use recollection to avoid this incorrect recognition (e.g., Dosher, 1984; Gronlund & Ratcliff, 1989; Hintzman & Curran, 1994; Jacoby, 1999; McElree, Dolan, & Jacoby, 1999; Rotello & Heit, 2000).

Response-speed methods were combined with the Remember-Know process-estimate method (see Section 1.5.2) by Gardiner, Ramponi, and Richardson-Klavehn (1999). In this study, participants were required to make Remember-Know responses after making recognition decisions at short (500 ms) or long (1500 ms) response deadlines. While the Remember-Know responses themselves were not speeded, more Know and Remember responses were made following the longer deadline. Gardiner et al. (1999) concluded that Know responses do not index an automatic familiarity process as suggested by some dual-process models, as both Remember and Know responses increased when the longer deadline permitted more effortful, controlled retrieval.

Further exceptions to the finding that familiarity is faster than recollection come from unspeeded Remember-Know studies where recognition decision RTs are analysed by which category of subjective experience the item was assigned to. Using this analysis many studies have shown recognition decisions for Remember responses to be reliably faster than for Know responses (e.g., Dewhurst & Conway, 1994; Dewhurst, Hitch, & Barry, 1998, Dewhurst, Holmes, Brandt, & Dean, 2006; Henson, Rugg, Shallice, Josephs, & Dolan, 1999; Vilberg & Rugg, 2007; Wixted & Stretch, 2004). However, it has been suggested that these patterns of findings occur because participants in Remember-Know experiments are instructed to respond Know only if the item is 'familiar and not recollected', i.e., they are required to wait until both processes have been completed before making a Know response (Yonelinas, 2002). In the current thesis, RTs were recorded for recognition decisions categorised by their later subjective experience judgment (Chapters 3 and 4) and in a novel analysis, RTs for post-recognition subjective experience, source, and confidence judgments were also measured (Chapter 4). The time course of recollection and familiarity processes is returned to in these chapters and discussed in terms of what the timing of responses can elucidate about how people make and understand judgments of subjective experience.

1.5.1.2. Associative versus single-item recognition

Dual-process models posit that recollection involves retrieval of qualitative information experienced when the item was encoded (Jacoby, 1991; Mandler, 1980; Tulving, 1985, Yonelinas, 1994, 2002). In associative recognition tests, where participants are asked to make judgments about some aspect of experimenter-controlled encoding such as whether two words were paired together for study, recollection should therefore aid retrieval of this associated information. On the other hand, familiarity or memory strength should be able to discriminate studied from non-studied items, but it should be less useful in tests of associative recognition. This is particularly true in 'pure' associative paradigms such as those used extensively by Naveh-Benjamin and colleagues (Kilb & Naveh-Benjamin, 2007; Naveh-Benjamin, 2000; Old & Naveh-Benjamin, 2008). Here items are studied in pairs and then participants are tested using both item and associative recognition tests. In the associative test, pairs of items are presented for test either in intact or recombined pairs and the participant is required to give an 'Old' recognition response only to intact pairs. Familiarity is of no help to this type of associative judgment as all the items on the test are familiar; recollection is required to determine whether a pair is intact or recombined. Performance on this type of associative test can therefore be used to index recollection, and can be compared against performance on item recognition for the same set of stimuli as item recognition is considered to involve both recollection and familiarity (Yonelinas, 2002). Item recognition and associative recognition can also be compared using a source monitoring procedure.

In source monitoring participants are asked to judge some aspect of the context in which an item was originally presented, or from what source they learnt the item. For example this could be whether an item was written in a particular font/colour/size, was shown on the left or right of the screen, was spoken in a particular voice, or a combination of these contexts (e.g., Meiser & Bröder, 2002). Participants' accuracy at retrieving this associative information is considered to index recollection (e.g., Wixted & Mickes, 2010), while their performance at identifying the item as old or new, regardless of whether their source judgment is correct, is taken as indexing both recollection and familiarity.

An advantage of comparing associative and item performance over comparison of recall and recognition is that test items can be the same across both retrieval conditions, the only change being the information requested or judgment required of the participant. However, as Yonelinas (1999, 2002) highlights, performance on some associative tests, such as recency judgments ('which list was this presented on, list 1 or list 2?'), may be influenced by familiarity as more recently studied items may be perceived as more familiar. In addition, the extent to which familiarity can contribute to associative recognition performance has been found to depend on whether the paired items or the item and the contextual information form a unified whole, with familiarity making a greater contribution to source memory if source and item information are unitized during encoding (Diana, Yonelinas, & Ranganath, 2008).

Task-dissociation methods are of particular importance to experiments presented in Chapters 3 and 4 of this thesis. Across the experiments presented in Chapter 3, three different type of recognition test are used: item recognition, associative 2-alternate-forced-choice (2AFC) recognition, and non-associative 2-alternate-forced-choice (2AFC) recognition. Within each of these test paradigms, experimental familiarity was manipulated in different ways and analysis examines how this influenced patterns of subjective experience. In Chapter 4, two experiments include measures of source memory and this objective measure of retrieval of context is compared to judgments concerning confidence and subjective experience.

1.5.2. Process-estimate methods

In contrast to task-dissociation methods where performance on different tasks is compared and inferences are made about the relative contributions of recollection and familiarity, process-estimate methods aim to provide quantitative estimates of the contribution of recollection and familiarity to overall recognition performance on a single task. There are three main process-estimate methods that have been used to examine recollection and familiarity parameters: the process-dissociation procedure, receiver-operating characteristic curves, and the Remember-Know paradigm.

1.5.2.1. Process-Dissociation-Procedure

The process-dissociation-procedure (PDP) is an extension to associative recognition procedures, discussed above. The PDP was developed by Jacoby (1991) in conjunction with his dual-process model of memory. In the PDP, participants study two separate lists of items and are tested under both inclusion and exclusion conditions. Under inclusion conditions, participants are asked to recognise which items they had studied, regardless of whether they had been studied on List 1 or List 2. Under exclusion conditions, participants are asked to only respond ‘yes’ if they studied the item on a particular list, List 1 for example. Performance in the inclusion condition is taken to reflect both familiarity and recollection processes, as both may contribute towards accurate recognition. In contrast, performance in the exclusion condition is thought to reflect recollection, as recollection is required to determine which list the item was studied on. In the PDP, estimates of recollection and familiarity are therefore calculated by comparing inclusion and exclusion performance. The contribution of recollection to performance is estimated by subtracting proportion correct in the exclusion condition from proportion correct in the inclusion condition: $R = P(\text{Inclusion}) - P(\text{Exclusion})$. The contribution of familiarity to performance is calculated by dividing performance in the exclusion condition by 1 minus the probability of recollection: $F = P(\text{Exclusion})/(1 - R)$.

One potential problem with using the PDP to measure recollection and familiarity is that it involves a rather strict criterion for the measurement of recollection (Yonelinas, 2002). The only form of recollection which 'counts' in this paradigm is recollection which allows the individual to determine which list the item was shown on. Noncritical recollection, for example recalling that you pictured your favourite coffee shop when the word 'café' was shown, is not helpful unless it supports list discrimination. Unless you recall that you formed the image when 'café' was seen on List 1, as opposed to heard on List 2 (if encoding was manipulated in this way), or you recall that your image of the coffee shop was one of the earlier images you formed (and therefore it must have been on List 1), recollection of the visual image is not going to help you discriminate whether 'café' should be reported or not. Studies have demonstrated that this type of noncritical recollection can influence estimates of recollection and familiarity, particularly when the two study lists are purposely very similar and participants are able to recollect other details from the study lists which do not aid list discrimination (e.g., Gruppuso, Lindsay, & Kelly, 1997; Mulligan & Hirshman, 1997; Wagner, Gabrieli, & Verfaellie, 1997; Yonelinas & Jacoby, 1996). However, other studies have concluded that the incidence of this occurring is low in standard PDP paradigms (Yonelinas, 2001a, 2001b, 2002; Yonelinas & Jacoby, 1996).

Yonelinas (2002) concludes that findings from the PDP converge with those of ROCs and the Remember-Know paradigm, which are detailed in the following sections, however, the similarity between the underlying assumptions of PDP and Remember-Know have been questioned. Gardiner, Java, and Richardson-Klavehn (1996) highlight that a major difference lies in that the Remember-Know procedure provides a first-person account of subjective experience, whereas the PDP account is third-person; and furthermore, that they differ in the way they define consciousness. For example, in their own comparison of these procedures Jacoby, Yonelinas, and Jennings (1997) state that by focusing on subjective reports, the Remember-Know procedure "…identifies consciousness with awareness. The inclusion/exclusion procedure, in contrast, defines consciousness with reference to intentional control of responding" (p. 41). Gardiner et al. (1996) and Richardson-Klavehn, Gardiner, and Java (1996) conclude that it is more sensible to regard the Remember-Know paradigm and the PDP as different, but complementary, ways of exploring recollection and familiarity.

1.5.2.2. Receiver-Operating Characteristic curves

Receiver-operating characteristic (ROC) curves can be plotted on recognition data where responses are made at differing levels of confidence, e.g., *6 – 'sure it was old'* to *1 –' sure it was new'*. ROC curves plot correctly recognised targets (hits) against incorrectly recognised lures (false alarms) at each level of confidence or decision criterion. The first point on an ROC curve includes only the items at the highest level of confidence, i.e., the items eliciting a *6*. The second point on the ROC includes both items eliciting the highest level of confidence and the items eliciting the next highest confidence, i.e., all the items assigned a *5* or a *6*. The shape of the ROC curve, the intercept with the y-axis, the area under the curve, and the slope of the zROC can then be calculated and compared across conditions. An example ROC and a zROC are plotted in Figure 1.2. Chance performance would be evidenced if the ROC was aligned on the diagonal (i.e., when hits are equal to false alarms), and the closer the ROC function gets to the upper left of the probability space, and the resulting greater area under the curve, the better the memory sensitivity or discriminability. A zROC is then used to quantify the shape of an ROC. The z-score of each hit and false alarm rate is calculated and plotted in z-space to produce a zROC. If the zROC is linear, then where it intercepts with the y-axis is taken as an estimate of recognition accuracy and the slope reflects the asymmetry of the ROC (Parks & Yonelinas, 2008).

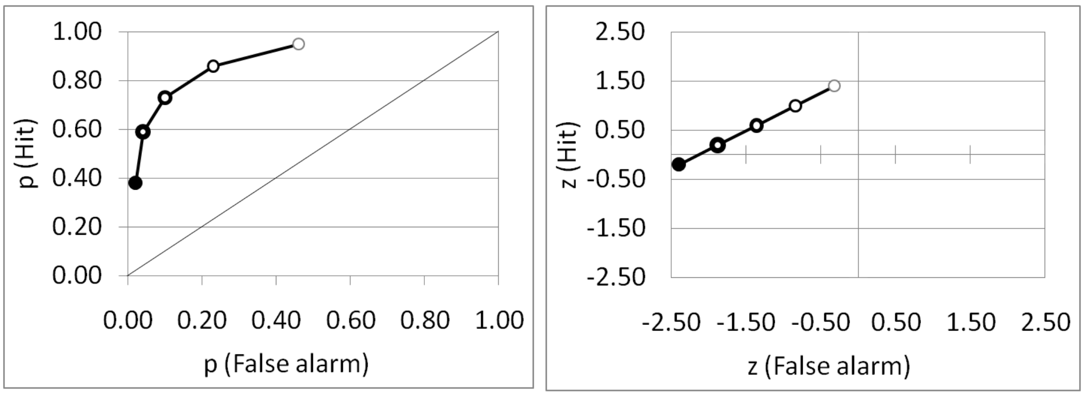

Figure 1.2. Example ROCs in probability space (left panel) and z-space (right panel) after Parks and Yonelinas (2008). The points in both panels represent pairs of hits and false alarms that are summed across a confidence scale ranging from 1 to 6, with 1 labelled 'Sure New' and 6 labelled 'Sure Old.' Thus, the first point at the farthest left represents hits and false alarms for items given a confidence rating of 6. The next point represents hits and false alarms for items given a confidence rating of 5 and 6, and so on down the confidence scale. Nb., the final point (for 1 or 'Sure New') is not plotted because it is constrained to be (1,1).

The analysis of ROC curves stems from signal detection theory (SDT) which assumes that a single process of familiarity or strength underlies recognition memory (see Section 1.4.3).

However, ROCs are now involved in the majority of models of recognition processes and the different models have different predictions regarding the shape of the ROC and the zROC (see Parks & Yonelinas, 2008; Yonelinas & Parks, 2007; for reviews). For example, because the EVSD model assumes that old and new distributions have the same shape, the model produces a symmetrical ROC that has a slope of 1.0 in z-space. In contrast, the UVSD model predicts asymmetrical ROCs with a slope in z-space of greater or less than 1.0 because the variance in the distribution of old items can differ from that of new items. Regarding estimates of recollection and familiarity, these two models assume that Remember responses simply reflect high-confidence recognition (e.g., Donaldson, 1996; Hirshman & Master, 1997; Wixted & Stretch, 2004) and predict that Remember-Know data should fall on the same ROC curve that is observed in standard ROC recognition studies (Parks & Yonelinas, 2008).

In contrast, the DPSD model predicts different ROC and zROC shapes depending on the relative contributions of recollection and familiarity to the type of recognition being tested. In item recognition, the model predicts curved asymmetrical ROCs that are pushed up on the left side due to the contribution of recollection to the symmetrical curve arising from the familiarity component. Additionally, the predicted item ROCs are approximately linear in z-space, however the threshold recollection process in the DPSD model leads the zROCs to be slightly U shaped (Glanzer et al., 1999; Parks & Yonelinas, 2008). For Remember-Know judgments, this model again predicts that these data should map onto the same ROC curve as confidence (Parks & Yonelinas, 2008).

In their reviews of ROC data from a wide range of recognition memory paradigms Yonelinas & Parks (2007; Parks & Yonelinas, 2008) concluded that single-component models of recognition memory do not adequately explain recognition and that there must be at least two processes underlying recognition. ROCs are at the centre of arguments regarding single- and dual-process models, but ROC analysis of data in the present thesis are not presented, as the focus was on manipulations of subjective awareness responses in different conditions and this was easier to interpret using measures of probability and ANOVA.

1.5.2.3. Remember-Know

As discussed in Section 1.4.2.3 above, Tulving (1985) introduced the Remember-Know paradigm into the memory literature by asking his participants to report their state of awareness for each item as either Remembered or Known. Over the last 26 years the paradigm has been used to demonstrate a wide variety of dissociative effects of experimental manipulations on Remember and Know states or dissociative effects between subjective states when comparing different populations of participants. As the Remember-Know paradigm forms the basis for all the experiments presented in this thesis detailed discussion of this process-estimate method is undertaken in the following sections.

1.6. The Remember-Know paradigm

1.6.1. Experimental dissociations

A wealth of studies have found evidence for experimental dissociations between Remembering and Knowing and four different patterns of results have been observed (for detailed reviews see Gardiner & Richardson-Klavenh, 2000, and Gardiner, 2008). Firstly, some experimental manipulations have been found to increase the proportion of Remember responses while not affecting Know responses. These manipulations typically include variables that engage more elaborative or conceptual processing at study. For example, increases in Remember responses have been found for deep versus shallow levels of processing, generation or vocalization versus reading of items at study (Gardiner, 1988; Gregg & Gardiner, 1991), orthographically distinctive versus orthographically common words (Rajaram, 1998), repetition of items at study (Dewhurst & Anderson, 1999), and for intentional versus incidental learning (Macken & Hampson, 1993). In addition, serial position effects have been shown to occur in Remembering but not in Knowing, and recall tests enhance later Remembering but do not improve Knowing (Jones & Roediger, 1995).

Secondly, the opposite pattern, that Know responses increase while Remember responses are unaffected, has also been demonstrated, particularly in relation to increases in perceptual processing at study. For example, presenting identical (Rajaram, 1993) or associatively related (Rajaram & Geraci, 2000) primes prior to presentation of a test word, implicit activation resulting from similar words being presented in a preceding lexical decision task (Dewhurst & Hitch, 1997), and suppression of focal attention (Mäntylä & Raudsepp, 1996) have all been shown to selectively enhance Know responses.

Thirdly, some manipulations lead to increases in Know responses while Remember responses are reduced. For example, non-word versus word recognition (Gardiner & Java, 1990), massed versus spaced repetitions of study list items (Parkin & Russo, 1993), and encoding faces with respect to their similarity versus encoding them with respect to their distinctiveness (Mäntylä, 1997) have all been found to elicit more Know responses and fewer Remember responses. Finally, the fourth pattern of results obtained through experimental manipulations is a parallel increase in both Remember and Know responses. This pattern has been demonstrated by manipulating response deadlines in recognition tests; both Remember and Know responses are increased in unspeeded compared to speeded recognition tests (Gardiner et al., 1999).

These different patterns of Remember-Know findings demonstrate what is termed 'functional independence' between Remembering and Knowing (Gardiner, 2008; Gardiner & Conway, 1999). Some variables affect Remembering but not Knowing; some variables affect Knowing but not Remembering; other variables have opposite effects on Remembering and Knowing; and yet other variables have similar effects on Remembering and Knowing. These functional dissociations have been taken as support for dual-process accounts by proponents of that viewpoint, however single-process protagonists such as Dunn (2004, 2008) have argued that the patterns observed can be explained using SDT models.

1.6.2. Special populations

In addition to experimental manipulations demonstrating dissociations between Remember and Know responses, dissociations are also evident in patterns of responding produced by different populations of participants. As mentioned in Section 1.3, reductions in recollection are considered to be one of the fundamental changes associated with decreased episodic memory performance in normal aging. This has been supported by experimental findings of older adults producing fewer Remember responses than younger adults (Friedman & Trott, 2000; Java, 1996; Norman & Schacter, 1997; Parkin & Walter, 1992; Perfect & Dasgupta, 1997; Perfect et al., 1995; Souchay et al., 2007).

For example, Parkin and Walter (1992) and Experiment 1 of Perfect et al. (1995) found that younger adults assigned approximately two thirds of recognised items to the Remember category, compared to one-third for older adults. In these experiments, the reciprocal one third and two thirds of items respectively were accompanied by familiarity; therefore, aging was associated with both decreases in recollection and corresponding increases in familiarity (Perfect & Dasgupta, 1997). However, a different pattern of responding was demonstrated by Mäntylä (1993) who found that older adults showed less recollective experience than younger adults but level of familiarity did not differ between age groups. In Mäntylä's task, recall cues had been self-generated at study and it was suggested that the stereotypic cues generated by older adults led to weaker encoding which then corresponded to a reduction of recollective experience at retrieval.

This possibility of an encoding deficit leading to reductions in recollective experience was explored by Perfect and Dasgupta (1997) using a verbal 'think aloud' protocol at encoding. Older adults reported less recollective experience and were shown to differ from young adults in how they encoded study material. Older adults were more likely to report failing to come up with an encoding strategy, and were less likely to use elaborative encoding when the stimuli were non-words. Further analysis demonstrated that once the difference in encoding strategies was accounted for, there no longer remained any age differences in recollective experience. This link between quality of encoding and subjective experience at retrieval has recently been explored using the metacognitive measure of judgments-of-learning (JOLs) made for items during study. Daniels, Toth, and Hertzog (2009; and also Souchay, Williams, Moulin, Isingrini, & Conway, 2006) instructed younger and older adults to predict the likelihood that they would retrieve an item later during study and then judge their retrieval as recollected or familiar at test. A strong association was found between prediction and level of recollection for younger adults, but the relationship was found to be much weaker for older adults, consistent with age-related deficits in recollection.

In sum, these studies demonstrate that conscious awareness concerning both encoding and retrieval of items is implicated in age-related declines in memory performance. Furthermore, different patterns of findings regarding whether a decrease in Remembering in aging is offset by an increase in Knowing, which can be used to support recognition performance, depends on how encoding strategies are controlled (Gardiner, 2008).

Remembering has also been shown to be much reduced in conditions such as Alzheimer's disease, amnesia, schizophrenia, and autism spectrum disorders. In these populations, while levels of Remembering are reduced, levels of Knowing are generally found to be unaffected or are affected to a much lesser extent than Remembering (Bowler et al., 2000; Dalla Barba, 1993, 1997; Huron et al., 1995; Knowlton & Squire, 1995; Schacter et al., 1997; Tanweer et al., 2009). As for older adults, the reduced levels of Remembering shown in these other populations are sometimes found to be offset by an increase in Knowing, resulting in no overall differences in memory performance between the population under study and the control group, e.g., in adults with Asperger's syndrome (Bowler et al., 2000); whereas other comparisons have found reduced levels of both Remembering and Knowing (e.g., amnesic patients; Knowlton & Squire, 1995).

The dissociations observed in these special populations using the Remember-Know paradigm are linked to the underlying changes in neurological functioning and/or anatomy for these groups. For example, the recollection deficits observed in aging have been linked to a decrease in frontal lobe functioning. Age-related neuronal loss has been found to be particularly pronounced in the frontal lobes (e.g., Haug & Eggers, 1991), and measures of frontal dysfunction such as performance on card sorting tasks and verbal fluency tasks have been shown to correlate with deficits in associative memory in aging (e.g., Craik, Morris, Morris, & Loewen, 1990; Fabiani & Friedman, 1996; Spencer & Raz, 1994) as well as with age-related decreases in Remember responses (Parkin & Walter, 1992; but see Perfect & Dasgupta, 1997). Further evidence linking behavioural responses to neurological findings are discussed in the following section.

1.6.3. Evidence from neuropsychology

Neuropsychological studies have demonstrated that patterns of recollection and familiarity impairment are related to patterns of brain damage in special populations, and functional neuroimaging has shown that different distributions of activation occur for different subjective experiences in normal functioning. Brief reviews of both are provided here; for comprehensive recent reviews see Eichenbaum, Yonelinas, and Ranganath (2007), Skinner and Fernandes (2007), and Wais (2008). Comparing recall and recognition performance in patients with amnesia, brain damage that includes both the hippocampus and surrounding temporal lobe has been found to disrupt both recollection and familiarity. However, such brain damage has been found to be more disruptive for recollection, as indicated by more severe deficits for recall performance over recognition performance (e.g., Hirst, Johnson, Phelps, & Volpe, 1988; Isaac & Mayes, 1999; Johnson & Kim, 1985; Volpe, Holtzman, & Hirst, 1986). Furthermore, studies which have used Remember-Know, PDP, or ROC measures have shown that when temporal lobe damage is extensive, recollection is disrupted to a much greater extent than is familiarity (e.g., Blaxton & Theodore, 1997; Knowlton & Squire, 1995; Schacter, Verfaellie, & Pradere, 1996; Schacter, Verfaellie, & Anes, 1997; Verfaellie & Treadwell, 1993).

If damage to the hippocampus is somewhat selective, some comparisons of recall and recognition performance suggest that recollection seems to be disrupted whereas familiarity is relatively spared (e.g., Aggleton & Shaw, 1996; Aggleton et al., 2000, 2005; Baddeley, Vargha-Khadem, & Mishkin, 2001; Bowles et al., 2010; Hanley, Davies, Downes, & Mayes, 1994; Holdstock, et al., 2000; Jäger et al., 2009; Mayes, Holdstock, Isaac, Hunkin, & Roberts, 2002; Mayes, et al., 2001; Turriziani, Serra, Fadda, Caltagirone, & Carlesimo, 2008; Vargha-Khadem et al., 1997; Yonelinas et al., 2002); however, others have reported impairments in both recollection and familiarity (Cipolotti, et al., 2006; Gold et al., 2006; Manns, Hopkins, Reed, Kitchener, & Squire, 2003; Wais, Wixted, Hopkins, & Squire, 2006). Differences in lesion extent and documentation, patient selection, the type of memory paradigm utilised, and overall memory impairment across studies have been identified as important factors to take into consideration in regard to the discrepancies in the literature (Bowles, et al., 2010).

Distinct patterns of event-related potentials (ERPs) have also been associated with different subjective experiences. In general, correct responses to Old items elicit more positive-going ERPs than correctly rejected new items, termed the ‘Old-New’ effect (see Friedman & Johnson, 2000; Rugg & Allan, 2000, for reviews). This effect begins approximately 300 ms after the onset of the stimulus and may last for several hundreds of milliseconds. Within this effect, several important subcomponents have been identified. For example, an early mid-frontal Old-New effect (300-500 ms), often called the “FN400”, which is thought to reflect familiarity; and a slightly later effect (400-800 ms) which is maximal over (left) parietal regions and is thought to index recollection (Curran & Doyle, 2011; Johansson, Mecklinger, & Treese, 2004; and see reviews by Curran, Tepe, & Piatt, 2006; Friedman & Johnson, 2000; Mecklinger, 2000, 2006; Rugg & Curran, 2007;). This later parietal effect, also referred to as the late positive component (LPC), has been associated with a variety of behavioural responses indicating recollection including both accurate source recognition (Senkfor & Van Petten, 1998; Wilding & Rugg, 1996), and Remember responses (Curran, 2004; Düzel, Yonelinas, Mangun, Heinze, & Tulving, 1997; Rugg, Schloerscheidt, & Mark, 1998; Smith, 1993; Trott, Friedman, Ritter, Fabiani, & Snodgrass, 1999).

Positron Emission Tomography (PET) and functional Magnetic Resonance Imaging (fMRI) can locate activity related to different recognition experiences. Indexing recollection by looking at encoding has found activity in the prefrontal, parahippocampal, and fusiform regions to be greater for items that later go on to be assigned to Remember rather than to Know (Wagner, Koutstaal, & Schacter, 1999; Henson, et al., 1999). Activations throughout the medial temporal lobe (MTL), including the hippocampus and parahippocampal gyrus, are found to predict subsequent free recall (Strange, Otten, Josephs, Rugg, & Dolan, 2002). Furthermore, hippocampal regions have been shown to be more active during encoding of pairs of items than during encoding of single items (e.g., Henke, Weber, Kneifel, Wieser, & Buck, 1999), and hippocampal activity has been found to be correlated with successful performance on source memory tests, thought to depend on recollection, whereas perirhinal cortex (PRc) activity has been shown to be correlated with item familiarity (e.g., Davachi, Mitchell, & Wagner, 2003; Kensinger & Schacter, 2006; Ranganath, et al., 2004; Tendolkar et al., 2008; Weis et al., 2004).

Similar patterns of activity have been observed at retrieval also. Bilateral hippocampal and parahippocampal regions have been shown to be related to associative recognition, which is more reliant on recollection, but not to item recognition, to which recollection and familiarity processes contribute (Yonelinas, Hopfinger, Buonocore, Kroll, & Baynes, 2001). In addition, left hippocampal and parahippocampal activation has been associated with Remember as opposed to Know responses (Eldridge, Knowlton, Furmanski, Bookheimer, & Engel, 2000; Henson et al., 1999). In the prefrontal cortex, left anterior prefrontal regions have also been found to be more active for Remembered items compared to Known items (Henson et al., 1999), though the same pattern was not found by Eldridge et al. (2000). Left anterior prefrontal regions have also been shown to be preferentially active in associative recognition compared to item recognition tests (Nolde, Johnson, & D'Esposito, 1998; Ranganath, Johnson, & D'Esposito, 2000; Raye, Johnson, Michell, Nolde, & D'Esposito, 2000). These findings suggest that left anterior prefrontal regions are involved in discriminations based on recollection to a greater extent than those based on familiarity.

In their review of the neuroimaging literature examining recollection and familiarity processes, Eichenbaum et al. (2007) conclude that the hippocampus is critical for recollection but not familiarity and that the parahippocampal cortex also contributes to recollection. In contrast, the perirhinal cortex is principally involved in familiarity-based recognition.

However, Wais (2008) questioned this conclusion and proposed that it resulted from the use of dual-process theory to interpret the results. Wais suggested that instead of indexing recollection and familiarity processes, the behavioural methods employed in the reviewed studies (e.g., Remember-Know) actually differentiated strong from weak memories. Wais (2008) performed a meta-analysis on 17 fMRI studies and demonstrated the same patterns of activation discussed above. However, he posited that when interpreted in terms of reflecting strong and weak memory rather than recollection and familiarity, the results do not suggest that the hippocampus or adjacent MTL structures are related to particular recognition processes (see also Squire, Wixted, & Clark, 2007, and Wais et al., 2006). In contrast to this viewpoint, Diana, Yonelinas, Ranganath (2010) have recently demonstrated that activation in the perirhinal cortex was related to successful retrieval of item details, whereas activation in the hippocampus and parahippocampal cortex was associated with recollection-based source retrieval. They concluded that this pattern of results is consistent with the idea that different MTL regions process different types of episodic information, and cannot be accounted for by memory strength.

These recent reviews and experimental findings demonstrate that while many areas have been demonstrated to be involved in recognition memory, the neural substrates of recollection and familiarity are not yet clearly delineated. Critically, interpretations of neuroimaging data are tied up within the pervading dual- versus single-process debate. Thus, a greater understanding of the basis of subjective report and the factors which influence Remember and Know responses is likely to be critical for the interpretation of neuroimaging studies.

1.6.4. Comparisons with confidence judgments

One type of manipulation that is particularly interesting for the dual- versus single-process debate is when patterns of responses for Remember-Know judgments and confidence judgments are compared. As discussed above, single-process viewpoints suggest that recognition of items depends on the underlying strength of the memory trace or confidence in memory and that conceptualising two underlying processes does not add to our understanding of recognition processes (e.g., Donaldson, 1996; Dunn, 2004). However, experiments that have compared Remember-Know and confidence judgments have shown different patterns of responding elicited by these two judgment types.

In early studies comparing confidence and Remember-Know, confidence was operationalised as a two-category scale of Sure-Unsure and proportion of items assigned to these two categories were compared against proportion assigned to Remember and Know. Gardiner and Java (1990) compared subjective experience judgments and confidence judgments to words and non-words in two separate experiments and demonstrated that these two types of judgment did not respond equivalently to word and non-word recognition. More Remember judgments were assigned to words compared to non-words while more Know judgments were made to non-words compared to words. In comparison, this interaction between word type and response was not found when Sure-Unsure judgments were made instead of Remember-Know judgments. Rajaram (1993) observed similar patterns comparing Sure-Unsure and Remember-Know judgments using a masked-priming manipulation. This manipulation led to increases in Know judgments while Remember judgments were unaffected. However, in a separate experiment where participants made Sure-Unsure judgments these differences were not found and levels of Unsure judgments did not differ across manipulations of perceptual fluency. These experiments were replicated by Tunney and Fernie (2007) with the addition of a Guess category for both subjective experience and confidence judgments. While patterns of responses were somewhat different across the priming manipulation, patterns of subjective experience and confidence responses again demonstrated that these two judgment types are not equivalent.

In a later set of experiments, Rajaram, Hamilton, and Bolton (2002) also replicated Gardiner and Java's (1990) word/non-word findings using a within-subjects design instead of a between-subjects design and in addition performed a between-subjects replication comparing amnesic patients and matched controls. This experiment demonstrated that while control participants produced the same patterns of results as participants in the within-subjects experiment, patterns of responses differed for amnesic patients. When making confidence judgments, patients' responses matched those of controls as they made more Sure than Unsure responses to both words and non-words. However, for subjective experience judgments patients did not show the crossover effect that words received more Remember responses while non-words received more Know responses. Amnesic patients' responses were evenly split across Remember and Know categories. Taken together, these findings demonstrate that while judgments of confidence and subjective experience may be interrelated, Remember-Know judgments are not made solely on the basis of confidence and the two judgment types are not "experimentally interchangeable" (Rajaram et al., 2002, p. 234).

Other experiments have compared Remember-Know and confidence judgments using a larger scale to measure confidence and have analysed them using ROC curves. Using this methodology, higher confidence has been consistently found to be associated with Remember responses, compared to Know responses (Rotello et al., 2004; Rotello, Macmillan, Reeder, & Wong, 2005; Slotnick, 2010; Wixted & Stretch, 2004; Yonelinas, 2001a; Yonelinas et al., 1996). Although these results are interesting for the single- versus dual-process debate, these studies have different numbers of confidence levels and subjective experience categories and therefore participants' response distributions for the two cannot be directly compared as was done when confidence was operationalised as Sure-Unsure.

Results from experiments comparing subjective experience and confidence judgments have been analysed, meta-analysed, and been subjected to different modelling approaches in order to determine whether they can be explained by single-process models (e.g., Donaldson, 1996; Dunn, 2004, 2008; Gardiner, et al., 2002; Macmillan, Rotello, Verde, 2005). Gardiner (2008) highlights that this debate involves many technical arguments concerning the appropriateness of the various different assumptions and statistical measures available for use in signal detection models. The current thesis does not aim to resolve this debate. When confidence and subjective experience are compared in experiments presented in Chapter 4, these judgments are taken at face value and conceptualised as reflecting the underlying awareness available to the participant for that item. Critically, the experiments in Chapter 4 also allow the subjective measures of confidence and subjective experience to be compared against source accuracy as a more objective measure of recollection.

1.6.5. Measurement issues within the Remember-Know paradigm

In the 26 years since Tulving (1985) introduced the Remember-Know paradigm researchers have tried and tested a number of variations on the paradigm and how best to calculate estimates of recollection and familiarity. Procedural questions include whether recognition and Remember-Know are best performed using one-step or two-step procedures and if this influences performance. A further question regards whether the original two categories of Remembering and Knowing are sufficient or whether other states of awareness can be identified. These measurement and procedural issues are discussed in the following sections.

1.6.5.1. One-step or two-step?

Hicks and Marsh (1999), Eldridge, Sarfatti, and Knowlton (2002), and Bruno and Rutherford (2010) have all examined whether recognition accuracy and patterns of responding differ if one-step or two-step recognition procedures are used. Two-step procedures involve separate judgments of recognition (Old-New) and subjective experience (Remember-Know(-Guess) – see Section 1.6.5.2 regarding inclusion of a Guess option). At test, participants initially judge whether they recognise an item, and then only if they do recognise the item are they asked to judge their state of awareness regarding that recognition. One-step procedures combine the recognition judgment with the judgment of subjective experience. Participants are asked to judge whether an item is Remembered, Known, or New. As discussed by Hicks and Marsh (1999), psychologically it might be easier for participants to weigh the evidence separately for recognition and subjective experience in a two-step procedure as opposed to having to consider all three subjective states simultaneously in a one-step procedure. Thus performance accuracy might be better for two-step procedures as opposed to one-step. Hicks and Marsh (1999) were also interested in whether asking participants to judge subjective experience in either one-step or two-step procedures influenced their recognition performance. Perhaps asking participants to think about their subjective experience underlying recognition of each item might increase their metacognitive awareness and thereby their recognition performance? Alternatively, it could increase the difficulty of the task for participants and reduce performance accuracy?

Hicks and Marsh (1999) compared recognition memory performance on a conventional Old-New recognition memory test with performance on using a one-step and two-step procedures. They examined hit rates, false alarm rates, SDT accuracy measures (*d′* and *A′*) and response criterion (*c* and B''_D indexes). Their results demonstrated that hit and false alarm rates with a conventional Old-New recognition test and a two-step Remember-Know procedure were equivalent, demonstrating that including Remember and Know as post-recognition judgments did not influence memory performance. However, hit and false alarm rates in the Old-New test and the two-step procedure were lower than with the one-step procedure. Additionally, Hicks and Marsh (1999) found that the one-step procedure provided more hits judged Remember than the two-step procedure. However, they also found that the one-step procedure also provided more false alarms judged as Remember and Know than the two-step procedure. Hicks and Marsh suggested that the more liberal response criterion observed in the one-step procedure was due to the greater difficulty of assessing all three states of awareness simultaneously in this task. Overall however, when both hit and false alarm rates were taken into account, none of the recognition test procedures resulted in differences in accuracy.

Eldridge et al. (2002) obtained similar results to those of Hicks and Marsh (1999). Fewer Know false alarms were obtained using the two-step procedure than using the one-step procedure and Know accuracy (*A′*) did not differ from chance using the one-step procedure. However, Bruno and Rutherford (2010) recently criticised the accuracy statistics employed in these two studies (*d′* and *A′*) and suggested that measuring recognition accuracy using d_a is preferable because it as a truly unbiased measure of recognition accuracy (Macmillan & Creelman, 2005).

Bruno and Rutherford's (2010) results demonstrated no differences in overall recognition accuracy between one-step and two-step procedures using either *d′* or d_a measures.

However, in line with the findings of Eldridge et al. (2002) and Hicks and Marsh (1999), differences were obtained for Know hits and false alarms using the different procedures. A greater proportion of items were assigned to Know as both hits and false alarms in the one-step procedure than in the two-step procedure. In addition, overall hit and false alarm rates were found to be higher in the one-step procedure than in the two-step procedure. This suggests that participants were more liberal in their response criterion in the one-step procedure, as was found by Hicks and Marsh (1999).

These findings suggest that when overall recognition accuracy is the primary measure of interest no differences are obtained through use of a one-step procedure in comparison to a two-step procedure. However, if you are more interested in the experiential states underlying recognition and how participants are using Remember and Know response options in different experimental conditions, these two methods have been shown to produce different patterns of responding, primarily for assignment of hits and false alarms to Know. In addition, if the focus is in measuring reaction time (RT) for recognition judgments separation of the recognition decision from the judgment of subjective experience, as in a two-step procedure, is critical for isolating the time taken to perform these separate processes. Reaction times for recognition decisions are analysed categorised by their subsequent post-recognition judgment in experiments presented in Chapters 3 and 4 of this thesis. Moreover, in a novel analysis in Chapter 4, RTs for post-recognition judgments were also measured. Analysis of recognition decision and judgment RTs are examined in terms of what the timing of responses can reveal about how people understand and make judgments of subjective experience.

1.6.5.2. Permitting guessing

In addition to comparing results obtained using one-step and two-step procedures, Eldridge et al. (2002) and Bruno and Rutherford (2010) also compared performance when participants were permitted to guess. A Guess category was first introduced to the Remember-Know paradigm by Mäntylä (1993), but papers by Gardiner and colleagues were the first that specifically set out to compare results from Remember-Know paradigms that did and did not allow Guess responses (Gardiner, Java, & Richardson-Klavehn, 1996; Gardiner, Kaminska, Dixon, & Java, 1996; and Gardiner, Richardson-Klavehn, & Ramponi, 1997).

Gardiner, Java et al. (1996) identified that inconsistencies in false alarm rates for Know responses in previous studies exploring levels-of-processing (LOP) effects may have been due to differences in the extent to which participants reported responses as Know when in fact they were guessing (cf. Rajaram, 1993; Jacoby et al., 1997). Gardiner, Java et al. (1996) allowed participants to report Guesses in their LOP experiment and demonstrated that accuracy of Guess responses was no different from chance. For Remember and Know responses they observed similar patterns of responding as obtained in a prior study by Gardiner (1988) where participants had been strongly discouraged from guessing and a Guess option was not provided. They concluded that participants sometimes use Know responses as a substitute for guesses when guessing is neither deliberately prohibited nor allowed, but when a Guess response option is provided Know responses demonstrate memory for the experimental episode whereas Guesses show no discriminative power.

Moreover, Gardiner et al. (1997) demonstrated that effects of a response criteria manipulation on patterns of Know responses by Strack and Förster (1995) were removed when a Guess option was permitted. Strack and Förster (1995) had concluded that patterns of Know responding were influenced by factors other than memory for the study episode; for example, judgmental strategies relating to response rates. In contrast, Gardiner et al. (1997) demonstrated that only Guess responses were affected by such factors.

This conclusion was supported by a later study that examined participants' reports of their conscious experiences underlying Remember, Know, and Guess responses. Gardiner, Ramponi, and Richardson-Klavehn (1998) found that justifications for Guesses showed evidence of various inferences and other judgmental strategies which were not directly related to the individual's memory for a studied word. In particular, Guess justifications often reflected familiarity with the items from some source other than the study episode, e.g., "*Holiday: I am eager to go on holiday so I am not sure whether I saw it here or whether I was thinking about it*"; or strategic responding based on inferences regarding the type of words that were on the studied list, e.g., "*Harp: It seemed that there were quite a few musical instruments, so I took a guess that it came up*" (both p. 8). Importantly, Know justifications did not include any evidence of use of inferences or judgmental strategies. In the present thesis, the justifications reported by Gardiner et al.'s participants were used as stimuli for all four experiments reported in Chapter 2, which examines layperson's understanding of experiential states and their experimental terminology. Additional findings from Gardiner et al. (1998) such as differences between Remember and Know justifications are therefore discussed in detail in that chapter.

The issue of whether inclusion of a Guess option alters patterns of responding was also examined by Eldridge et al. (2002) and Bruno and Rutherford (2010) in their comparisons of one-step and two-step procedures. Eldridge et al. (2002) demonstrated that including a Guess response in a one-step procedure increased accuracy of Know responses due to the removal of experimental ‘noise’ caused by participants assigning guesses to the Know category. This also led to a reduction of the false alarm rate for Know responses and an improvement in overall accuracy (*A’*). Bruno and Rutherford (2010) also found a reduction in false alarm rate for Know responses when a Guess option was included in a one-step procedure, however this was not accompanied by an improvement in overall accuracy when measured by d_a instead of *A’*. Bruno and Rutherford (2010) concluded that, using unbiased statistical comparisons, accuracy is not influenced by whether participants are permitted to Guess. However, as in the one-step versus two-step comparison above, if your primary interest is reports of subjective experience as opposed to overall memory accuracy, inclusion of a Guess category is beneficial as it reduces the likelihood that participants will assign guesses to the Know category (Gardiner, 2008).

1.6.5.3. Fractionation of Knowing and Familiarity

Whilst the addition of a Guess response option is now pretty standard in Remember-Know experiments, how Know responses are conceptualised varies greatly across studies. In their review of the Remember-Know literature in 2000, Gardiner and Richardson-Klavehn identified interpretation of Know responses as “the most vexatious problem in the remember/know paradigm” (p. 238). The root of this problem lies in, firstly, whether Know responses are defined to participants in terms of familiarity or certainty; and relatedly, whether Know responses are interpreted as reflecting an underlying process of familiarity or a state of knowing.

A number of researchers have acknowledged this problem and have attempted to clarify how they defined and interpreted Know responses in their experiments. For example, some researchers choose to ask participants to make Remember-Familiar judgments instead of Remember-Know. Dobbins, Kroll, and Liu (1998, p. 1309) state “We chose to use the word 'familiar' because students often confuse the more standard 'know' response with an expression of high confidence". Similarly, Donaldson, MacKenzie, and Underhill (1996) assert "…*familiar* rather than *know* was used to indicate nonrecollection, because the word *know* carries a connotation of certainty that is inconsistent with a confidence rating that indicates lack of certainty. Participants find it hard to say that they are unsure that an item was there but that they know it was." (p. 487, italics in original). Other researchers choose to encompass both familiarity and knowing within one response category, for example, Kelley and Jacoby (1998) define Knowing as "…the inability to recollect any details of the study presentation in combination with *a feeling of familiarity or certainty* that the word was studied" (p. 134, italics added). The issue of how Know responses are defined and interpreted is not helped by the fact that many research reports, including those by Dobbins, Kroll, and Liu (1998) and Donaldson et al. (1996) above, do not publish the exact wording that was used to define response categories to participants.

To overcome the problem of the concepts of familiarity and knowing being encompassed in one response option Conway et al. (1997) added a separate 'Familiar' category to their participants' response options. Conway et al.'s (1997) experiment was a longitudinal study of student learning. In piloting their materials they found that for some multiple-choice questions (MCQs) students reported that their subjective state for an answer was one of neither recollection or familiarity, instead they felt that they 'just knew' the answer. In their conceptualisation of this subjective state, Conway et al. draw attention to two different bases of familiarity. Typically familiarity is considered to arise from a feeling that an item has been encountered recently, but without any recollection of the encounter. This feeling of familiarity leads participants to believe that the item was one they studied. However, Conway and colleagues point out that familiarity can also arise from having encountered an item frequently, as opposed to recently and uniquely. In this case an item may come to mind without recollection, but also without the feeling of a recent encounter which cannot be identified. Here, the item has become semanticised in memory and the subjective state accompanying recognition is one of 'knowing'. This is akin to the original conceptualisation of Remembering and Knowing reflecting retrieval from episodic and semantic memory systems respectively (Tulving, 1985).

In Conway et al.'s (1997) study, students took MCQ exams following four psychology lecture courses and one course, Introduction to Psychology, was later retested. At initial testing higher performing students were found to designate more answers as Remembered than did poorer performing students, however at re-test this pattern had reversed and higher performing students assigned more answers to Know than they did to Remember. Conway et al. discuss this 'R-to-K shift' as reflecting a change in knowledge representation from episodic to semantic memory brought about by loss of episodic details from memory and the emergence of conceptual organisation. Students initially performed better when they could recollect some element of the learning episode, but six months later students performed better when their knowledge of psychology concepts had become more semantic in nature.

Evidence of the R-to-K shift has been demonstrated in similar studies of student learning by Barber, Rajaram, and Marsh (2008) and Herbert and Burt (2001, 2003, 2004). Herbert and Burt (2003) examined the effects of different review opportunities on the R-to-K shift, comparing combinations of MCQs and short-answer questions (SA). Participants who had been given the most varied opportunity for re-study (MCQ plus SA) showed greater evidence of an R-to-K shift. They concluded that schematisation of knowledge was facilitated when students underwent regular testing as this gave them the opportunity to review their knowledge; and enhanced when a variety of test formats were used. The conclusion that the R-to-K shift reflected a schematisation of knowledge which enhanced performance was supported by results from a separate assessment of quality of student learning included by Herbert and Burt (2001, 2003) in their studies.

In a study exploring learning of rare words definitions, Dewhurst, Conway, and Brandt (2009) also found a shift from Remembering to Knowing over time. Participants were tested at delays of five minutes, four weeks, eight weeks, and six months, and the proportion of Remember responses was found to dominate on the first test and decrease on subsequent tests, while Know responses increased. Dewhurst et al. (2009) also found that Remember and Know responses did not differ in accuracy or confidence, though both were significantly more accurate and associated with higher levels of confidence than Familiar and Guess responses. This further validates the separation of the Know and Familiar categories of subjective awareness. Taken together, these findings demonstrate that people are able to consciously appreciate the differences between the experiential states of Remember, Know, Familiar, Guess; and that, for a range of materials, the subjective experiences associated with learning demonstrate consistent patterns over time.

It could be argued that the separation of Know and Familiar might only apply to learning paradigms with rich materials such as used by Conway et al. (1997) and Herbert and Burt (2001, 2003, 2004). In typical recognition memory experiments, which use lists of words as the to-be-studied material and recognition is only measured at one time point, participants are not able to integrate or semanticise the studied information into any body of knowledge and therefore the experiential state of 'Knowing' might not appear to be applicable. However, Dewhurst et al. (2009) were interested in whether the R-to-K shift would be observed for less 'meaningful' materials and asked participants to learn rare words and their definitions specifically because these were unrelated facts that would be less easy to integrate into a wider schema. While these materials are arguably more meaningful than lists of words, they are still less meaningful than the academic material used in previous studies. In Dewhurst et al.'s (2009) study, participants assigned 10% of items to Know the first time they were tested, a situation which is analogous to a single-time-point recognition experiment. While this proportion is a lot lower than the 22% assigned to Know in Conway et al.'s (1997) study, it demonstrates that even for less meaningful learning, participants considered that the Know response reflected their experiential state for some items only five minutes after study and with no opportunity for integration into a wider body of knowledge[2]. These 10% of unrelated facts were 'just known' without any recollection or feelings of familiarity regarding the study episode only a short time after said study episode. This suggests that the separation of Know and Familiar might also be applicable to the subjective experiences associated with learning of materials in more typical recognition paradigms such as words, names, and paired-associates where the materials have lower intrinsic meaning. A key theme of the current thesis is to examine whether the separated concepts of Knowing and Familiarity have any utility in standard episodic recognition memory paradigms, and indeed whether there may be differences between Know and Familiar in the same way that there are differences

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

between Remember and Know. To this end, all four subjective experience response categories are employed in all experiments.

1.6.5.4. Other variations on Remember-Know categories

Other researchers, particularly those interested in the neural correlates of recollection and familiarity have modified the Remember-Know categories in other ways. Using complex scenes as stimuli, Montaldi, Spencer, Roberts, and Mayes (2006) trained participants to distinguish between recollection and three levels of familiarity: very weak familiarity, moderate familiarity, and strong familiarity. Using fMRI, Montaldi et al. demonstrated patterns of neural activation that increased or decreased linearly as familiarity increased from weak to strong. In addition, hippocampal activity was not modulated by changes in familiarity; the hippocampus was only activated for items that were recollected. Using a similar paradigm where participants were instructed to assign recollected items to Remember and non-recollected items to four levels of confidence, Yonelinas, Otten, Shaw, and Rugg (2005) also observed patterns of neural activity that increased or decreased linearly with increasing familiarity and that were differentiated from the neural activity associated with recollection.

In an eyewitness memory paradigm, Palmer et al. (2010) also adapted the typical Remember-Know judgment. Their procedure involved participants viewing a 3-minute video and later being asked to identify one male and one female ‘suspect’ from line-ups. Participants rated the confidence of their identification and were asked whether their identification was based on Remembering or Knowing. Additionally, participants were asked to justify their subjective experience judgment and these justifications were independently classified by trained raters as providing evidence of relevant recollection. *Relevant* recollection was that which contained clear and convincing evidence that the participant recollected contextual details from their previous exposure to the target in the video, e.g., “I remember seeing her outside the bank”. *Non-relevant* recollection included responses that indicated the absence of recollection, e.g., “she seemed familiar”, or uncertainty, e.g., “I remember a girl at a bank, but I’m not sure if it was her in the lineup” [*sic*]. In this strict classification procedure, responses that claimed the recollection of contextual detail but did not provide any specificity about the detail recalled, e.g., “I remember his face” or “I remember what she was doing in the video” were also classified as *non-relevant* recollection-based responses. Taking confidence ratings, Remember-Know judgments, and recollection ratings into account, results demonstrated that positive identifications accompanied by Remember judgments were more likely to be accurate than those accompanied by Know judgments; however, further analysis demonstrated that taking Remember-Know responses into account did not improve identification accuracy beyond that found by examining confidence ratings. Critically however, considering relevant and non-relevant recollection statements was found to improve identification accuracy beyond that achieved by examining confidence ratings.

While the procedures employed by Palmer et al. (2010), Montaldi et al. (2006), and Yonelinas et al. (2005) do not map directly onto the four states of subjective experience, Remember, Know, Familiar, Guess, used in the present thesis, they do demonstrate that other researchers are exploring alternatives to the dichotomous Remember-Know paradigm and the underlying theoretical suggestion that there are additional states of subjective awareness accessible for study. In particular, the Palmer et al. (2010) experiment demonstrates that participants are able to consciously access more information than is elucidated by a standard Remember-Know judgment and this further information can be diagnostic of memory accuracy. This is also being explored by proponents of the argument that recollection is a continuous process. For example, Wixted (2010; Wixted & Mickes, 2010) asked participants to make a Remember-Know judgment in conjunction with a 20-point confidence judgment and a judgment of source. His findings demonstrated that when only the highest confidence Remember and Know judgments were compared they were equally accurate in terms of recognition sensitivity, however they were accompanied by different levels of source accuracy. The relationship between confidence, accuracy, source, and recollective experience is the focus of experiments presented in Chapter 4. Critically, these experiments aimed to examine the relationship between source accuracy and the separated subjective experience categories of Know and Familiar, as well as Remember, and the overlap between objective and subjective reports concerning retrieval of contextual details.

1.7. Aims of this thesis

The central aim of this thesis was to explore how people make and understand judgments of subjective experience. Within this, the experiments presented in Chapters 2, 3 and 4 explore lay conceptions of subjective awareness, how participants overcome experimental familiarity, and the relationships between source, confidence, and subjective experience respectively.

The key themes of this thesis are: How should the subjective experiences of Know and Familiar be defined and understood in recognition memory? What is the relationship between confidence and subjective experience? And how do objective manipulations influence subjective experience? How each of these key themes is explored in Chapters 2, 3, and 4 is outlined below.

The first aim of the current experiments was to explore how Know and Familiar subjective experiences can be differentiated. This was examined in terms of lay interpretations based on reports of others' memory experiences (Chapter 2); how manipulations of familiarity (Chapter 3) and confidence (Chapter 2) influence Know and Familiar responding; and the accuracy, confidence, source accuracy, and speed with which people make Know and Familiar responses (Chapters 3 and 4). These experiments permit the examination of how the separate Know and Familiar categories of subjective experience are used in straightforward single-item episodic recognition memory tasks and whether there are differences between Know and Familiar in the same way that there are differences between Remember and Know.

The second aim of this thesis was to examine the relationship between confidence and subjective experience. Lay interpretations of others' memory experiences were explored in terms of confidence ratings and assignment of memory experience justifications to categories of subjective experience when confidence is manipulated (Chapter 2). The influence that confidence and subjective experience judgments have on each other when the two judgments are made consecutively post-recognition is then examined along with the relationship between these two subjective judgment types and source accuracy as an objective measure of retrieval of contextual detail (Chapter 4). Here the primary question of interest is how participants interpret and act upon the relationship between confidence and subjective experience.

The final aim of this thesis was to investigate how subjective experience responses are influenced by objective manipulations of confidence (Chapter 2), familiarity (Chapter 3), and source (Chapter 4). An additional methodological point of interest here is whether any influence of objective manipulations is dependent on whether the manipulation was carried out within- or between-subjects (see Section 3.1.5). The influence of objective manipulations on subjective experience responses is crucial to theoretical interpretations of how people understand their memory experiences.

While findings from these experiments may add to support for dual-process and/or single-process accounts of recognition, and are discussed in relation to these theoretical accounts where appropriate, this thesis did not set out to provide evidence in support of one particular viewpoint. Rather, the primary focus of the experiments presented here is the conscious awareness that people are able to access during recognition. In this thesis, reports of experiential state are taken at face value and analysis is interested in how, why, and when, participants retrieve items with particular subjective awareness.

2. USING LAY CONCEPTIONS OF SUBJECTIVE EXPERIENCE TO EXPLORE REMEMBERING AND KNOWING, AND FINDING FAMILIAR

"I did not rely much on the confidence ratings when I made 'Remember' judgements. However, I tended to rely more on the confidence ratings when I was hesitating between one of the three other categories, in the cases where subject's [sic] *justifications were more fuzzy."*
(Participant 83, Experiment 2.3)

This chapter examines the relationship between subjective experience and confidence of recognition by exploring how lay-people interpret others' memory experiences. Using three internet-based questionnaires, the relationship between subjective experience and confidence was examined (Experiment 2.1) and manipulated (Experiments 2.2 and 2.3) to assess to what extent people understand memory experiences in terms of confidence and how this interacts with use of the subjective experience categories of Remember, Know, Familiar, and Guess. Using a laboratory task, Experiment 2.4 explored the legitimacy of separating the classic Know category of subjective experience into separate Know and Familiar categories.

2.1. GENERAL INTRODUCTION

As described in Chapter 1 (Section 1.4.4), advocates of a single-process viewpoint such as Dunn (2004, 2008) posit that instead of Remember and Know responses reflecting two different underlying memory processes or subjective experiences they reflect different strengths on a single scale of memory trace or confidence. This contentious relationship between Remember-Know judgments and confidence has been around since the formulation of the Remember-Know paradigm by Tulving (1985) who, in the very first use of the paradigm, demonstrated that Remember responses were associated with higher confidence ratings than Know responses. This has been explored and replicated in a variety of experimental paradigms (e.g., Brewer & Sampaio, 2006; Gardiner & Java, 1990; Rajaram, 1993; Rajaram et al., 2002; Rotello & Zeng, 2008; and Yonelinas, 2001a). Comparison of experimental confidence and subjective experience judgments is the focus of Chapter 4.

Although the relationship between subjective experience and confidence has been acknowledged from the conception of the paradigm, and has been the subject of much debate between single-process and dual-process theorists, how the lay-person understands, conceptualises, and acts upon subjective experiences of memory and confidence in memory are still not understood.

An obstacle to gaining insight into lay-peoples' understanding of subjective experiences of memory is that researchers have not been consistent in how they define categories of subjective experience. Previous research has shown that even subtle differences in Remember-Know procedures can lead to significant differences in how people assign responses to Remember and Know (e.g., Bastin & Van der Linden, 2003; Bodner & Lindsay, 2003; Gardiner et al., 1997; Geraci & McCabe, 2006; Kihlstrom, Kim & Dabady, 1996; McCabe & Balota, 2007; Norman & Schacter, 1997; Rotello et al., 2005; Strack & Förster, 1995). As stressed by McCabe and Geraci (2009), it is of critical importance to ensure that Remember-Know instructions are consistent across research studies due to the ongoing theoretical debates around interpretation of Remember-Know judgments and inferences based on meta-analyses and aggregated data (e.g. Dunn, 2004; Gardiner et al., 2002; Parks & Yonelinas, 2007; Wixted & Squire, 2004).

In all experiments in the current thesis participants are asked to assign responses to one of four categories of subjective experience: Remember, Know, Familiar, and Guess. As discussed in the General Introduction (Section 1.6.5.3) these four categories of subjective experience were first delineated by Conway et al. (1997) and have subsequently been used by Barber et al. (2008), Dewhurst et al. (2009), and Herbert and Burt (2001, 2003, 2004). They could also be considered somewhat similar to the four categories utilised in the neuroimaging work of Montaldi et al. (2006) and the work of Kihlstrom et al. (1996) and Brewer and Sampaio (2006). As explained in Section 1.6.5.3, the central reason for splitting the classic Know category into separate Know and Familiar categories were problems of definition. Geraci, McCabe and Guillory (2009) recently informally collated experimental definitions of Remembering and Knowing from other researchers and found large differences in the amount of detail provided to experimental participants, whether instructions emphasised recollection should be specifically of something from the study episode, and whether Knowing was defined as a high-confidence state of certainty or a low-confidence state based on a feeling of familiarity. Kelley and Jacoby (1998) highlight the problem of definition by describing Knowing as ". . .the inability to recollect any details of the study presentation in combination with *a feeling of familiarity or certainty* that the word was studied" (p. 134, italics added).

2.1.1. The influence of terminology and confidence on Remember-Know judgments

Geraci et al. (2009), McCabe and Geraci (2009), and Rotello et al. (2005) recently explored how experimental definitions of subjective experience categories influenced participants' Remember-Know judgments in a variety of experimental manipulations. In two experiments, Rotello et al. (2005) compared traditional Remember instructions (after Rajaram, 1993) with more conservative Remember instructions. These conservative instructions specified that participants should only respond Remember if they could actually describe specific details of the study episode and that they might need to justify their responses to the experimenter. Results demonstrated that, under conservative instructions 83% of Remember responses were made at the highest level of confidence; however when instructions were less conservative only 65% of Remember responses were made at the highest level of confidence, with 16% made at the second highest level (with the remaining 19% assigned to the three lower levels of confidence).

This influence of source-specific Remember instructions on Remember-Know judgments was explored more recently by McCabe and Geraci (2009), who demonstrated that instructions that aimed to constrain recollection to the study episode led to reduced Remember hits and false alarms and increased Know hits and false alarms. Replicating Rotello et al., participants became more conservative in their use of the Remember response. In a second experiment, McCabe and Geraci explored how the use of the neutral terms of 'Type A' and 'Type B' instead of Remember and Know influenced judgments. Results demonstrated that use of neutral terminology reduced the number of Remember false alarms for both younger and older adults thereby increasing memory accuracy. McCabe and Geraci highlighted that one of the problems with the use of the terms Remember and Know in memory experiments is that confusion may arise from the pre-existing connotations participants have for them. Lay-persons' understanding of 'remember' and 'know' may well not be in relation to distinct states of awareness associated with retrieval from memory. McCabe and Geraci concluded that to obtain the most accurate Remember-Know data from participants neutral terminology and source-specific instructions should be utilised.

In two further experiments, Geraci et al. (2009) explored whether Remember-Know responses reflected different subjective experiences or levels of confidence by emphasising Know responses as either highly confident or as less confident. Participants in Geraci et al.'s experiments underwent two study-test sessions one week apart. In both sessions they studied mixed lists of words and non-words, as this is a factor which has been shown to dissociate confidence judgments from Remember-Know judgments (Gardiner & Java, 1990; Rajaram et al., 2002; see General Introduction, Section 1.6.4). In the test phase of session one participants were asked to assign recognised items to Remember or Know, and in session two were asked to assign recognised items to the confidence categories of Sure and Unsure. When confidence was emphasised in the Know definition, results replicated the standard finding that Remembering and Knowing were differently influenced by words and non-words, whereas confidence responses were not. However, when confidence was not emphasised in the definition of Know responses the patterns of data were found to be similar for Remember-Know and Sure-Unsure responses. Geraci et al. (2009) concluded that the wording of instructions can have important theoretical implications for our understanding of Remembering, Knowing, and confidence. In all the experiments presented in this thesis, source-specific instructions are included in definitions of Remember, Know, Familiar, and Guess responses and Remember, Know, and Familiar definitions are each accompanied by a real-world example of the subjective experience, see Appendix A.

2.1.2. Participants' understanding of Remember-Know definitions

One further problem with the Remember-Know paradigm highlighted by Geraci et al. (2009) is the number of participants they found who did not understand how to apply the Remember-Know and confidence judgments. Geraci et al. asked participants to complete a post-test questionnaire to assess their understanding of experimental instructions. Responses demonstrated that approximately 20% of their participants did not understand the Remember-Know instructions and in one experiment a few participants did not understand the confidence instructions. Geraci et al. excluded these participants from analysis and highlighted how important it is to perform post-experiment manipulation checks to ensure that participants are using responses in the way intended. One common way to do this is to ask participants to justify their recognition responses, i.e., state *why* they said they recognised that item. Gardiner et al. (1997) used this form of manipulation check, subsequently analysed the content of the obtained justification statements, and published a full list of 270 justifications of Remember, Know, and Guess recognition decisions in Gardiner et al. (1998). These justifications were used as stimuli[3] in the four experiments presented in the current chapter and are shown in Appendix B.1.

In the initial Gardiner et al. (1997) experiment, after participants had completed the recognition test and assigned each recognised item to Remember, Know, or Guess, the experimenter chose at random two responses from each category and asked the participant to explain what it was that led them to recognise the word as one they had studied. Importantly, the emphasis was on what led to the recognition decision, participants were not asked to justify why they had assigned a word to a particular category of subjective experience. When analysing the transcripts of the justifications Gardiner et al. (1998) classified responses by their most salient characteristics. For Remember justifications two expert raters classified the justifications as involving: intra-list associations, extra-list associations, item-specific images, the item's physical features, and self-reference. The first four of these Gardiner et al. discussed as being related to participants' attempts to memorise the study list via use of effortful strategies, associations, and imagery. The final category of self-reference was suggested as reflecting items automatically triggering awareness of a personal memory from everyday life. Examples of the responses for each category are shown in Table 2.1.

[3] With permission from John Gardiner.

Table 2.1. Example Remember justification statements for each classification type, from Gardiner et al. (1998).

Cue Word	Justification	Classification
President	Yesterday I associated this word with the word "minister". Today I automatically remembered about that association.	Intra-list association
Kilt	I remembered that I thought of a Scottish man.	Extra-list association
Gun	When I saw it yesterday I had an image of a gun and I thought it was a strange word to put in.	Item-specific image
Sauerkraut	I remembered it because I could not pronounce it yesterday!	Item's physical features
Harp	On Friday I was in a restaurant with a harpist. I remember thinking of that yesterday.	Self-reference

In comparison to Remember justifications, for Know and Guess justifications Gardiner et al. (1998) noted that transcripts were shorter and used rather limited vocabulary. This was suggested to reflect the fact that participants found it harder to articulate the reasons behind their recognition decision and required more encouragement from the experimenter. Gardiner et al. indentified Know justifications as reflecting attributions of recent unremembered encounters as they were absent of recollection of specific contextual details and were instead characterised by feelings of familiarity, just knowing, thinking a word occurred, or reporting of the absence of recollective details. Examples of Know justifications are shown in Table 2.2. Also shown in Table 2.2 are examples of Guess justifications demonstrating the various inferences and other judgmental strategies employed by participants when their recognition decision was based on guessing. These show that Guess justifications were primarily speculations about recent encounters neither Remembered nor Known. Inferences were either to do with different types of words from the study list or knowledge inferred from information about appropriate response rates.

Table 2.2. Example Know and Guess justification statements from Gardiner et al. (1998).

Type	Cue Word	Justification
Know	Gun	I just knew that I knew it.
	Kilt	It seemed familiar but I wondered.
	Butterfly	It was one of those words that rang a bell.
	Gun	There was no association, I just had a feeling that I saw it, I was sure.
	Professor	It was not in any of the little stories I made up to remember the words, but I had a strong feeling of familiarity.
	Squirrel	I remembered something about squirrels, but I cannot remember what.
Guess	Officer	It was just a guess.
	Slipper	I saw this word somewhere recently at some point, but I am not sure whether it was there yesterday.
	Father	I kept saying "no" so I just guessed it was there because you said that 50% of the words were there.
	Harp	It seemed that there were quite a few musical instruments, so I took a guess that it came up.
	Holiday	I am eager to go on holiday so I am not sure whether I saw it here or whether I was thinking about it.

To assess the relationship between subjective experience and confidence expressed in people's memory justifications, the expert raters in Gardiner et al. (1998) coded which Know and Guess justifications indicated certainty or uncertainty about the accuracy of the recognition decision. This was done through counting how many responses included phrases such as 'sure'/'not sure', 'confident'/'not confident', or 'I know'/'I think'. For Know responses both raters rated 25% of justifications as indicating participants felt sure of their recognition, and between 11 and 20% as indicating uncertainty. For Guess justifications 72-77% were rated as uncertain and none were rated as certain.

Through their exploration of how their participants differently justified recognition decisions based on Remembering, Knowing, and Guessing, Gardiner et al. (1998) demonstrated that these three categories of subjective experience reflected access to different memory processes at study and retrieval. Furthermore, rating of justifications also revealed differences in the confidence levels associated with the different subjective experience categories.

2.1.3. Aims of the current experiments

The central aim of the four experiments presented here was to explore how lay-people understand reports of memory experiences in terms of confidence and categories of subjective experience. The novel approach used here is to not ask participants about their own memory experiences but to place participants in the role of 'memory expert' and ask them to classify other people's reports of recognition decisions. Three internet-based questionnaires and one laboratory task are reported here which all used the justification statements of Gardiner et al. (1997, 1998) as stimuli. Experiment 2.1 asked questionnaire respondents to assign a confidence rating to the justification statements. These confidence ratings were then used to select justifications for use in Experiments 2.2 and 2.3 where participants were asked to categorise justifications as Remember, Know, Familiar, or Guess when confidence was manipulated. Finally, in Experiment 2.4 participants were asked to split original Know justifications into two types (undefined by the experimenter) to examine whether the splitting of the Know response category into separate Know and Familiar categories fit with lay-persons' understanding of memory experiences and to see if there were any systematic, rational means by which the Know items could be split.

Previous experimental tasks have demonstrated that Remember responses are typically associated with higher confidence ratings than Know responses (Brewer & Sampaio, 2006; Gardiner & Java, 1990; Rajaram, 1993; Rajaram et al., 2002; Rotello & Zeng, 2008; Tulving, 1985; and Yonelinas, 2001a). Experiment 2.1 was designed to examine whether this experimental finding was replicated when it is not the participants' own contents of memory under study. The topic of exploration here was whether lay understanding of subjective experiences links reports of experiential state to confidence in the same way that experimental tasks have found. If people are able to interpret others' justifications in terms of the memory processes underlying them, and these are linked to confidence, then each category of justification statement should be assigned different confidence ratings in Experiment 2.1.

Experiments 2.2 and 2.3 were designed to manipulate the confidence associated with a report of subjective experience in a method somewhat akin to the experimental paradigms of Geraci et al. (2009), McCabe and Geraci (2009), and Rotello et al. (2005). In those experiments, the confidence associated with Remember-Know judgments was manipulated through the terminology used to define Know responses (Geraci et al., 2009), or instructions regarding how conservative Remember responses should be (McCabe & Geraci, 2009; and Rotello et al., 2005). In Experiment 2.2, confidence was manipulated through use of justification statements which obtained high, medium, and low confidence ratings in Experiment 2.1. In Experiment 2.3, confidence was manipulated by each justification statement being accompanied by a confidence value that was either appropriate or inappropriate to the subjective experience category of the justification. In both experiments, instead of giving a confidence rating for the item, here the task for participants was to assign these justifications to one of the four subjective experience categories. If participants' interpretations of the memory processes underlying the justifications are influenced by the confidence associated with the justification, then this should be reflected in how the justifications are assigned to the categories of subjective experience.

Experiment 2.4 focuses on the issue of whether the Know category of subjective experience can be reliably divided into separate categories of Know and Familiar. Previous research has shown reliable patterns of results for Know and Familiar responses when participants are given the four response options of Remember, Know, Familiar, and Guess in experimental tasks which demonstrates that people can differentiate these experiential states when exploring their own memory experiences (Barber et al., 2008; Conway et al., 1997; Dewhurst et al., 2009; and Herbert & Burt, 2001, 2003, 2004). Using Gardiner et al.'s (1998) justification statements again, Experiment 2.4 explores whether participants can differentiate the underlying experiential states of Know and Familiar from others' reports of their memory experiences. Based purely on intrinsic understanding of memory experiences, if participants are able to identify two different categories of subjective experience in the original Know justifications this would lend further support to the validity of the Know-Familiar separation.

2.1.4. Splitting Know into Know and Familiar

The problem of the unity of the concept of Knowing is exemplified by the fact that Gardiner et al.'s (1998) expert raters rated 25% of Know responses as indicating certainty and up to 20% as indicating uncertainty. This suggests that there was a great deal of variance within the subjective experiences underlying Know responses in their experiment.

One aim of the experiments in the current chapter was to examine whether some of the Know justifications provided by Gardiner et al.'s (1997, 1998) participants perhaps lent more towards being a justification of Familiarity rather than of Knowing. To address this, Chris Moulin and I independently categorised Gardiner et al.'s (1998) original 90 Know justifications as reflecting either a Familiar or Know recognition response based on the definitions of Know and Familiar shown in Appendix A. Comparing our independent ratings revealed that our categorisation had matched on 74 of the 90 items giving an inter-rater reliability of .82. For the remaining 16 items each was discussed until consensus was reached. Of the 90 items, final classification of expert ratings was that 47 were justifications reflecting a Know response and 43 were justifications based on Familiarity (see Appendix B.1 for which were classed as Know and which as Familiar). These separate Know and Familiar categories of justification statement were examined separately in all the experiments presented here.

2.2. Experiment 2.1: Lay understanding of subjective experiences in terms of confidence

2.2.1. Introduction

When subjective experience and confidence ratings are made by participants in a memory experiment Remember responses are typically associated with higher confidence ratings than Know responses (Brewer & Sampaio, 2006; Gardiner & Java, 1990; Rajaram, 1993; Rajaram et al., 2002; Rotello & Zeng, 2008; Tulving, 1985; and Yonelinas, 2001a). It is not surprising that subjective experience and confidence are related, as information from one's subjective awareness is likely used to judge how confident one is about a memory (Gardiner, 2001). However, the aim of the current experiment was to examine whether the differences in confidence associated with Remember-Know judgments are reflected in lay-persons' judgments of others' memory experiences. To this end, questionnaire respondents were asked to rate how confident they thought a previous participant had been in their recognition decision based on only their justification statement. If participants' confidence ratings are reliably different for justification statements reflecting the four types of subjective experience it would demonstrate that lay people are able to understand and interpret others' reports of memory experiences in a systematic way. Importantly, after Gardiner et al.'s (1998) original Know justifications have been split into separate Know and Familiar categories of justification, if reliable differences are found between confidence ratings to Know and Familiar justifications this will add support to the validity of using these two separate response categories in experiments.

2.2.2. Method

2.2.2.1. Participants

Data was collected using an online questionnaire which was active from April 2008 to September 2009. The questionnaire was advertised on a number of international psychology web experiment lists as well as being sent out to friends, family, and Leeds Memory Group (LMG) participant email lists. Some participants were University of Leeds Psychology undergraduates and were given one Participant Pool credit for participation. Full data sets from 309 participants were obtained (225 female, 84 male; mean age = 28.1, SD = 10.78, range = 16 to 65).

2.2.2.2. Materials and Design

The cue words and justification statements to be used as items in the online questionnaire were those published in the appendix of Gardiner et al. (1998). In their experiment, participants had been asked to justify why they thought they had recognised an item and each participant provided six justifications, two for Remember responses, two for Know responses, and two for Guess responses. For use in the current questionnaire all 270 items were sorted into 10 lists of 27 items for ease of completion by participants online. Each list contained equal numbers of Remember, Know, and Guess justifications (9 of each). Participants were randomly assigned to list when they accessed the web page and within each list items were presented in random order[4].

Some participants in Gardiner et al. (1997, 1998) had not used all response categories, therefore there were 8 missing items (7 x Guess, 1 x Know). In the lists where these missing items occurred they were replaced with items from other lists, chosen at random. To ensure that participants' responses to one justification were not influenced by another justification, any justifications that referred to other items were split onto separate lists or

[4] As participants were randomly assigned to which of the 10 lists of justifications they saw the exact number of responses per item was not equal. Each item received between 27 and 35 responses.

altered so as to no longer refer to another item. For example, for *policeman* "*This word followed the word gun so I know they were both there"*, here the items *gun* and *policeman* were split onto separate lists so that a participant would not see both the justification including *"gun"* and the cue word *gun*.

2.2.2.3. Procedure

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

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

2.2.3. Results

To examine whether confidence levels associated with memory justification statements differed, analysis was conducted in two ways. Firstly, the mean confidence ratings given to justifications of each subjective experience category were calculated. This was performed using an items-analysis whereby the mean confidence rating given to each of the 270 items was calculated and then these means became the data for further analysis. Secondly, analysis explored how usage of the confidence ratings differed across the subjective experience categories so the proportion of responses that were assigned to Remember, Know, or Guess justifications at each confidence level was calculated. These analyses were then repeated with the Know justifications split into Know and Familiar justifications based on the expert ratings (Section 2.1.4).

2.2.3.1. Confidence assigned to Remember, Know and Guess justifications

The mean confidence ratings made to items from each subjective experience category are shown in Table 2.3. Remember justifications received the highest confidence ratings, followed by Know and then Guess. ANOVA demonstrated a significant main effect of subjective experience (as a between-items factor) $F(2,267) = 424.38$, $p < .001$, and t-tests showed significant differences between confidence ratings assigned to each category of subjective experience (all $p < .001$).

Table 2.3. Mean confidence ratings assigned to Gardiner et al.'s (1998) Remember, Know, and Guess justifications (and standard deviations).

Subjective Experience	Mean Confidence Rating (%)
Remember	81.62 (7.21)
Know	52.56 (13.50)
Guess	34.34 (11.29)

2.2.3.2. Proportion of responses at each confidence level

To explore the patterns of usage of confidence ratings, the proportion of responses at each confidence level which were assigned to Remember, Know, and Guess were calculated and are shown in Figure 2.1. The proportion of Remember justifications assigned to each level of confidence remained below .12 until the 50% confidence level but then increased sharply with confidence level. The proportion of Know justifications assigned to each level of confidence was largest around the central levels of confidence, with between .38 and .50 of justifications that were assigned to confidence levels from 30 to 75% being Know justifications. The proportion of Guess justifications assigned to each level of confidence was found to decrease steadily as confidence level increased.

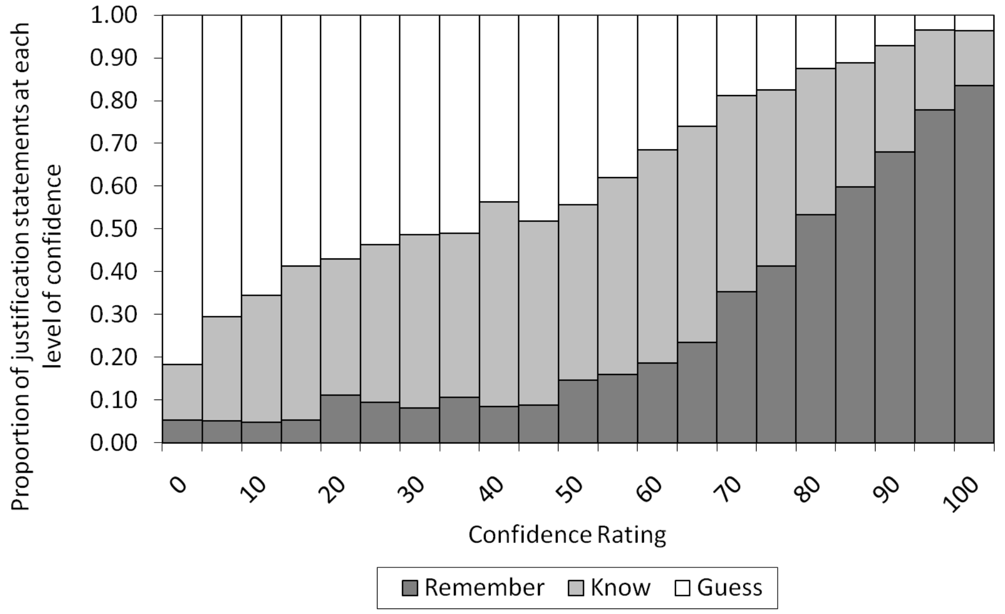

Figure 2.1. Proportion of Gardiner et al.'s (1998) Remember, Know, and Guess justification statements assigned to each level of confidence.

2.2.3.3. Confidence assigned to Remember, Know, Familiar, and Guess justifications

To examine whether the split of the Know justifications provided by Gardiner et al.'s (1998) participants into separate Know and Familiar categories (based on expert categorisation of the statements by HLW and CM, see Section 2.1.4) was reflected in the confidence values assigned to the justifications, the above analysis was repeated using the four categories of subjective experience. The mean confidence ratings made to items from each subjective experience category are shown in Table 2.4.

Table 2.4. Mean confidence ratings for Remember, Know, Familiar, and Guess justifications (and standard deviations)[5].

Subjective Experience	Mean Confidence Rating (%)	Number of Items
Remember	81.62 (7.21)	90
Know	58.18 (11.29)	47
Familiar	46.42 (9.47)	43
Guess	34.34 (11.29)	90

[5] The values for Remember and Guess are identical to those shown in Table 2.3 as these justifications types were not affected by the splitting of Know into Know and Familiar.

Remember justifications received the highest confidence ratings, followed by Know, then Familiar, then Guess. ANOVA again demonstrated a significant main effect of subjective experience (as a between-items factor), $F(3,266) = 321.43$, $p < .001$, and t-tests showed significant differences between confidence ratings assigned to each category of subjective experience (all $p < .001$). Different levels of confidence were reliably associated with the separate Know and Familiar categories of subjective experience.

2.2.3.4. Proportion of Know and Familiar responses at each confidence level

To explore whether the separation of the Know category of subjective experience into Know and Familiar was also reflected in the patterns of usage of confidence ratings, the proportion of original Know responses at each confidence level which were assigned to Know and Familiar were calculated and are shown in Figure 2.2. The pattern of results shows large fluctuations with between .30 and .50 of responses at lower levels of confidence being to Know justifications. However, a steady increase in proportion of justifications assigned to Know was evident at confidence ratings from 60% upwards.

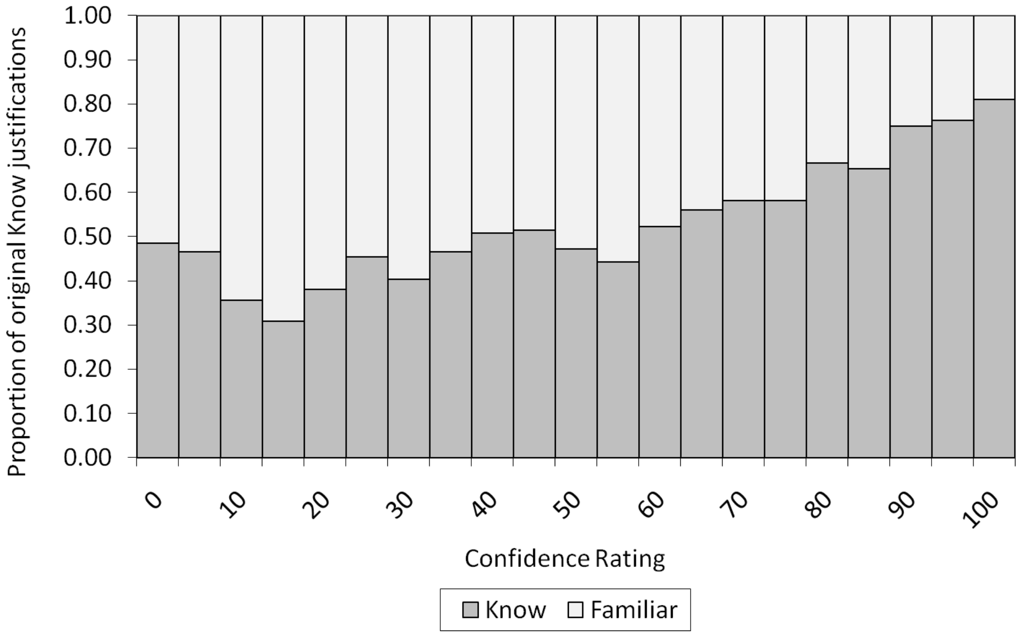

Figure 2.2. Proportion of original Know responses that were assigned to Know and Familiar at each confidence level.

2.2.4. Discussion

The focus of this discussion is to summarise the central findings of this experiment; more substantive theoretical discussion is reserved until the General Discussion of this chapter.

Results demonstrated that mean confidence assigned to justification statements originally made for Remember, Know, and Guess responses differed significantly. Remember justifications were assigned higher confidence ratings than Know justifications and in turn Know confidence ratings were higher than Guess confidence ratings. This suggests that not only is there a relationship between subjective experience and confidence when it the contents of one's own memory that is being assessed (Gardiner & Java, 1990; Dewhurst et al., 2009; Rajaram, 1993; Rajaram, et al., 2002; Rotello & Zeng, 2008; Tulving, 1985; and Yonelinas, 2001a), but that judgments of confidence based on the experiential reports of others show a reliable relationship also. Lay people understand the relationship between subjective experience and confidence in the same way as observed in experimental paradigms.

Critically, re-analysis of mean confidence values after fractionation of original Know justifications into Know and Familiar categories demonstrated significantly higher mean confidence values assigned to Know compared to Familiar justifications. This supports the experimental findings of Brewer and Sampaio (2006), Dewhurst et al. (2009) and Kihlstrom et al. (1996) that Know responses are associated with higher levels of confidence than Familiar responses. The current findings suggest that people are also able to differentiate these two experiential states based on others' memory reports and, at least in part, this discrimination is related to the level of confidence associated with the subjective experience. This finding adds validity to these two subjective states being provided as separate response options in experimental paradigms.

The patterns of usage of the confidence levels revealed that the relationship between confidence and subjective experience was not a direct one. Guess justifications had a near linear relationship with confidence, as the proportion of Guess justifications assigned decreased with confidence, whereas the proportion of Remember justifications assigned were small at low levels of confidence but then increased sharply above 50% confidence. A similar pattern was demonstrated by the fractionated Know and Familiar justifications with instability at lower levels of confidence but a steady increase in proportion of justifications assigned to Know at confidence ratings above 60%.

This experiment has demonstrated that different categories of subjective experience justifications receive different ratings of confidence. Although this experiment did not take a first-person approach to understanding subjective experience as it asked participants to rate others' memory descriptions, the findings do lend support to the idea that it is subjective awareness that gives rise to confidence (Gardiner, 2001; Tulving, 1985). However, this experiment only examined the relationship between subjective experience and confidence in one direction, what about whether different levels of confidence can be categorised as different types of subjective experience? This is the direction of the relationship that is explored in Experiments 2.2 and 2.3.

2.3. Experiment 2.2: Lay understanding of subjective experience in terms of Remember, Know, Familiar, and Guess judgments

2.3.1. Introduction

The aim of this experiment was to examine how manipulations of confidence influence lay-persons' understanding of subjective experiences of memory. In previous experimental memory tasks, Geraci et al. (2009) manipulated the confidence associated with Remember-Know judgments by changing the confidence implied in the definition of Knowing and McCabe and Geraci (2009) and Rotello et al. (2005) changed how conservative the instructions were for Remember responses. Both these manipulations resulted in different patterns of Remember-Know responses. The current experiment was interested in whether different patterns of subjective experience categorisation were found when judging memory justifications associated with different levels of confidence. To manipulate confidence in the current experiment, the confidence ratings obtained in Experiment 2.1 were used to select justification statements associated with high, medium, and low levels of confidence to be used as stimuli. The task for participants was to assign justifications to the Remember, Know, Familiar, and Guess categories of subjective experience.

This experimental approach assumes that if subjective experience and confidence are related, then it should be possible to manipulate them. The current experiment presents justifications that are associated with different levels of confidence and explores whether the original subjective experience underlying the justification can be recovered. If participants' categorisation of items is influenced by the confidence associated with a justification, results should show different patterns of categorisation at different levels of confidence. For example, a statement that originally justified a Remember response (in Gardiner et al., 1998) but which only received a low confidence rating (in Experiment 2.1) may not be consistently assigned to the Remember category. There may also be differences in how confidence affects the categorisation of the different types of justification statement. For example, the evocative recollections, intra- and extra-list associations, specific item information and self-reference associated with Remember responses may permit these justifications to remain unaffected by manipulations of confidence, but assignment of Know, Familiar, and Guess justifications to the appropriate category might be more greatly influenced by the confidence level associated with the item.

2.3.2. Method

2.3.2.1. Participants

Data was collected using an online questionnaire which was active between January and September 2009. Advertisement was undertaken in the same manner as in Experiment 2.1. Full data sets from 502 participants were obtained (388 female, 111 male, 3 no response; mean age = 26.71, SD = 12.40, range = 16 to 85).

2.3.2.2. Materials and Design

The cue words and justification statements to be used as items in the online questionnaire were a selection of those used in Experiment 2.1. Using the mean confidence ratings obtained from participants in Experiment 2.1, within each subjective experience category (Remember, Know, Familiar, and Guess) statements were sorted from highest to lowest confidence assigned. The two statements with the highest, medium[6], and lowest confidence values were selected from each category giving a set of 24 items to be used in this questionnaire[7]. For each participant items were presented in random order.

---

[6] Based on the median.

[7] Where cue words were duplicated across categories one of the items was exchanged for the neighbouring item e.g. *grasshopper* + justification was swapped for *body* + justification, confidence values 33.45 and 33.83 respectively, see Appendix B.1.

During data collection a second selection of 24 items was collated from items with the next lowest, medium, and highest confidence ratings to be used as a second version of the questionnaire. The two versions were then swapped online intermittently in order that approximately equal numbers of participants completed each. Full data sets were obtained from 248 participants for version 1 and from 254 participants for version 2. As no differences were observed between the data from the two versions all analyses were conducted on both versions together. The justification statements used in this experiment are marked in Appendix B.1 and the maximum, minimum, and mean confidence ratings for High, Medium, and Low confidence justifications for each subjective experience category are shown in Table 2.5.

Table 2.5. Maximum, minimum, and mean confidence ratings for the selected justifications from each subjective experience category.

Subjective experience	Confidence level	Max. confidence rating	Min. confidence rating	Mean confidence rating
	High	94.77	91.61	93.00
Remember	Medium	82.12	82.04	82.07
	Low	66.61	46.36	58.28
	High	88.57	79.09	84.11
Know	Medium	58.13	54.17	56.40
	Low	41.67	30.56	34.88
	High	67.58	57.50	61.96
Familiar	Medium	46.07	43.33	44.90
	Low	35.87	30.54	32.50
	High	62.59	55.00	57.66
Guess	Medium	34.40	33.13	33.87
	Low	13.67	7.22	10.16

As can be seen from Table 2.5, confidence ratings did overlap across subjective experience categories. For example, Remember justifications which had achieved the Lowest confidence ratings (mean = 58.28) had been assigned similar confidence ratings to both Medium confidence Know items (mean = 56.40) and High confidence Guess items (mean = 57.66). This is interesting as whilst these justifications were rated as reflecting the same levels of confidence in recognition in Experiment 2.1, they had originally justified difference subjective experience responses in Gardiner et al.'s (1997, 1998) experiment. The current experiment was interested in whether the original subjective experience categories could be 'recovered' from the justifications despite the differences in confidence within categories of subjective experience, and the similarities in confidence across the categories.

2.3.2.3. Procedure

The procedure for this experiment closely followed that of Experiment 2.1 though here instead of being asked to make a confidence rating, participants were presented with the cue word and justification statement and asked to classify that justification as a Remember, Know, Familiar, or Guess judgment. The instructions given at the start of the online questionnaire regarding how the justifications statements were obtained matched those given for Experiment 2.1. Participants were then instructed that in this experiment they would be shown the justification statements and that their task was to make a judgment about the person's recognition decision by classifying it as Remember, Know, Familiar, or Guess. Full definitions of the categories were provided (see Appendix A.1) and it was emphasised that participants should ensure that they fully understand the definitions before they left that web page as the definitions would not be shown again. It was suggested that participants could write down the definitions to refer to later. Participants were also told that they may only select one category for each statement but that they should try to use all the categories at least once, and to read all the justifications carefully as some of the justifications would be easier to categorise than others. Participants then entered their demographic information before being shown the first item. For each item, participants were shown the cue word and justification statement accompanied by the question "Which recognition category?" and the options Remember, Know, Familiar, and Guess. They made their decisions by clicking the appropriate button. After seeing all 24 items participants were debriefed as in Experiment 2.1.

2.3.3. Results

This experiment was concerned with whether participants were able to reliably match justification statements to their original category of subjective experience and whether categorisation would differ depending on the confidence level assigned to the statement by participants in Experiment 2.1. Analysis in this experiment took the form of items-analysis with the original subjective experience of the item (REMEMBER, KNOW, FAMILIAR, GUESS) and the confidence level of the item (High, Medium, Low) being treated as between-items factors and the subjective experience response of the participant (Remember, Know, Familiar, or Guess) as a within-items factor. For clarity, within the Results section only, when CAPITALISED the terms REMEMBER, KNOW, FAMILIAR, GUESS refer to the original subjective experience of the item and when only Initial Letter is capitalised they refer to the subjective experience response of the participant.

The proportion of items within an original subjective experience category and a particular level of confidence that were assigned to each of the subjective experience response categories were calculated across participants, e.g., the proportion of High confidence REMEMBER justifications that were assigned to Remember by participants. Firstly a 4(original subjective experience) x 3(confidence level) x 4(subjective experience response) ANOVA was performed. Although the proportions summed to 1 within each confidence level and therefore between-items factors of confidence level and original subjective experience could not be calculated, this form of analysis was favoured as it allows interactions between variables to be explored. Separate ANOVAs and further comparisons were performed to examine patterns within the different levels of confidence and original subjective experience categories.

The 4(original subjective experience) x 3(confidence level) x 4(subjective experience response) ANOVA demonstrated a significant main effect of subjective experience response, $F(2.03,73.12) = 9.11$, $p < .001$, and significant interactions between subjective experience response and original subjective experience, $F(6.09,73.12) = 41.12$, $p < .001$, and subjective experience response and confidence level, $F(4.06,73.12) = 14.13$, $p < .001$. There was also a significant three-way interaction between original subjective experience, confidence level, and subjective experience response, $F(12.19,73.12) = 5.55$, $p < .001$ (all Greenhouse-Geisser corrected). These interactions are now considered each in turn.

2.3.3.1. Analysis of two-way interactions

The means for the significant interaction between subjective experience response and original subjective experience are shown in Figure 2.3. When confidence level is not considered, the majority of REMEMBER and FAMILIAR justification statements were appropriately allocated to their original subjective experience category. Conversely, for KNOW justifications nearly as many were assigned to Familiar as to Know, and for GUESS statements a large proportion were assigned to Familiar instead of Guess.

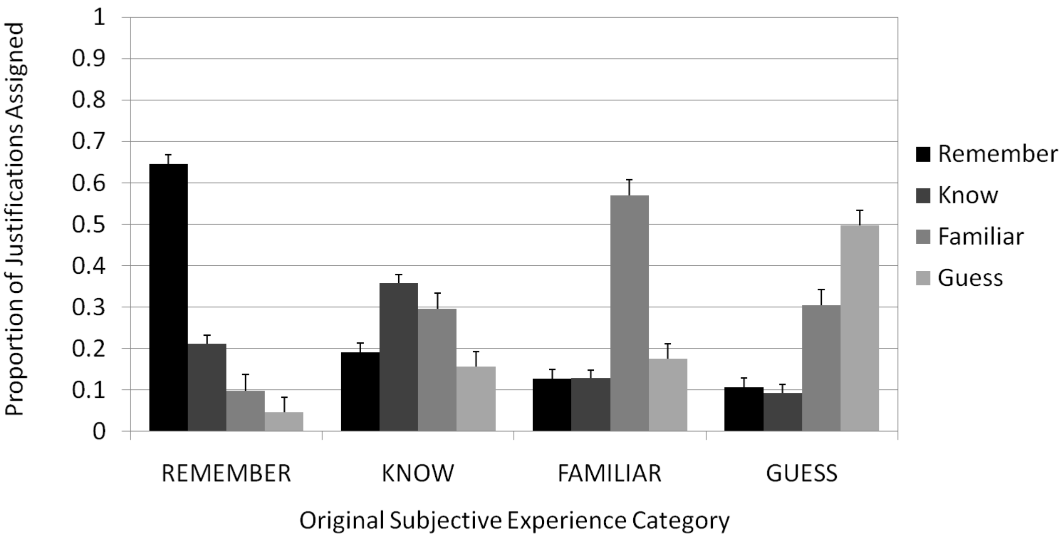

Figure 2.3. Proportion of justifications of the four original subjective experience categories that were categorised as Remember, Know, Familiar, and Guess. Errors bars = 1 SeM.

The interaction of confidence level with subjective experience response is explored in the following four sections of analysis. Firstly, within each of the original subjective experience categories a 3(confidence level) x 4(subjective experience response) ANOVA was carried out. If significant effects were found then separate ANOVAs were performed comparing participants' use of the four subjective experience responses at each of the three levels of confidence and comparing use of each of the four levels of subjective experience response across confidence levels. Results of these analyses are reported in the next four sections.

2.3.3.2. Assignment of High, Medium, and Low confidence REMEMBER justifications to Remember, Know, Familiar, and Guess

How confidence level interacted with subjective experience response for REMEMBER justifications is shown in Figure 2.4. For statements originally justifying a REMEMBER recognition decision in Gardiner et al.'s (1998) experiment, irrespective of it receiving a High, Medium, or Low, confidence rating by participants in Experiment 2.1, the majority of these REMEMBER justifications (between .57 and .68) were appropriately assigned to the Remember category of subjective experience. A 3(confidence level) x 4(subjective experience response) ANOVA demonstrated a significant main effect of subjective experience response, $F(3,27) = 209.50$, $p < .001$, and a significant interaction between subjective experience and confidence level, $F(3.75,16.86) = 3.64$, $p = .028$ (Greenhouse-Geisser corrected). Effect of confidence level was not calculated as proportions summed to 1. Separate ANOVAs at each level of confidence demonstrated significant main effects of subjective experience response at High, $F(3,9) = 97.46$, $p < .001$, Medium, $F(3,9) = 261.46$, $p < .001$, and Low, $F(3,9) = 25.16$, $p < .001$, confidence levels.

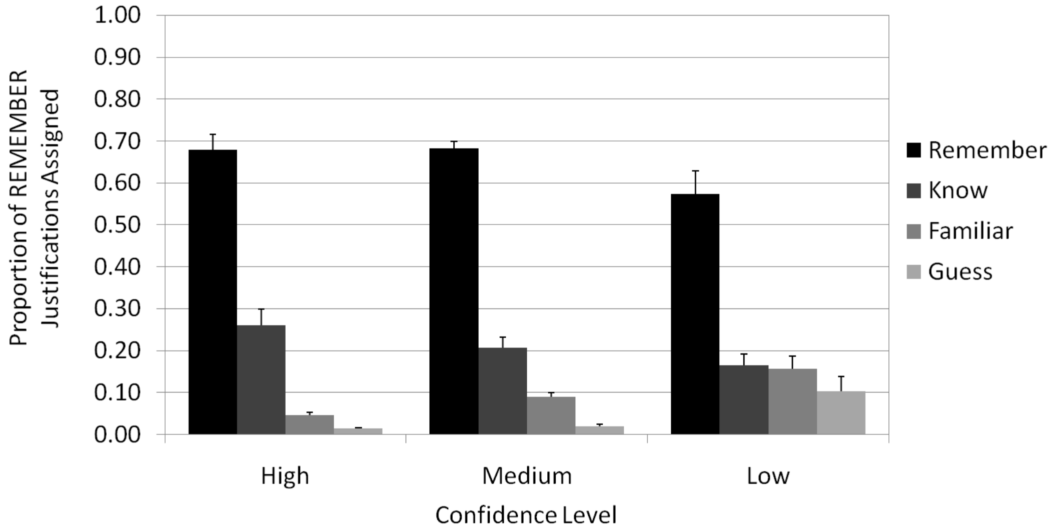

Figure 2.4. Proportion of REMEMBER justifications given High, Medium, and Low ratings of confidence in Experiment 2.1 that were assigned to Remember, Know, Familiar, and Guess. Error bars = 1 SeM.

Planned t-tests for High confidence justifications demonstrated that all comparisons of subjective experience response were significantly different, confirming that the majority of justifications were assigned to Remember, followed by Know, then Familiar, then Guess (all at least $p < .04$). The same result was found for Medium confidence responses (all at least $p < .04$). For Low confidence responses, more justifications were assigned to Remember than to Know, Familiar, or Guess (all at least $p < .02$), however the only other significant comparison was that significantly more justifications were assigned to Familiar than to Guess, $t(3) = 4.19$, $p = .025$.

ANOVAs comparing use of subjective experience response categories across confidence levels demonstrated that only for Familiar and Guess responses was a significant main effect of confidence level found: Familiar, $F(2,9) = 9.69$, $p = .006$, Guess, $F(2,9) = 6.36$, $p = .019$. Confidence level did not significantly influence the proportion of REMEMBER justifications assigned to Remember, $F(2,9) = 2.49$, $p = .14$, or Know, $F(2,9) = 2.33$, $p = .15$. Planned comparisons revealed that the likelihood of a justification being assigned to Familiar or Guess decreased as confidence level increased. A greater proportion of REMEMBER justifications were assigned to Familiar if they were of Medium confidence compared to High confidence, $t(6) = 3.96$, $p = .007$, and this comparison for Low compared to Medium levels of confidence approached significance, $t(6) = 2.21$, $p = .069$. In addition, a greater proportion of REMEMBER justifications were assigned to Guess if they were of Low confidence compared to Medium confidence, $t(6) = 2.44$, $p = .05$.

In sum, these results demonstrate that participants consistently assigned the majority of REMEMBER justifications to the Remember category, regardless of the confidence associated with the justification. In addition, the proportion of REMEMBER justifications that were assigned inappropriately to the Know category did not differ with confidence either, only the proportion of justifications assigned to Familiar and Guess were found to differ depending on the confidence level associated with the justification.

2.3.3.3. Assignment of High, Medium, and Low confidence KNOW justifications to Remember, Know, Familiar, and Guess

In contrast to the results for REMEMBER justifications shown in Figure 2.4, the patterns for assignment of KNOW justifications to subjective experience categories demonstrate that categorisation of KNOW justifications was strongly influenced by confidence level, see Figure 2.5. For High confidence justifications the majority of KNOW justifications (over.60) were appropriately assigned to Know, however for both Medium and Low confidence KNOW justifications the majority of justifications were assigned to Familiar (.36 and .44 respectively), followed by Know (.32) if confidence was Medium, and Guess (.31) if confidence was Low. A 3(confidence level) x 4(subjective experience response) ANOVA demonstrated a significant main effect of subjective experience response, $F(3,27) = 6.79$, $p = .001$, and a significant interaction between subjective experience and confidence level, $F(6,27) = 10.30$, $p < .001$. Separate ANOVAs at each level of confidence demonstrated significant main effects of subjective experience response at High, $F(3,9) = 16.26$, $p = .001$, Medium, $F(3,9) = 8.10$, $p = .006$, and Low, $F(3,9) = 4.39$, $p = .037$, confidence levels.

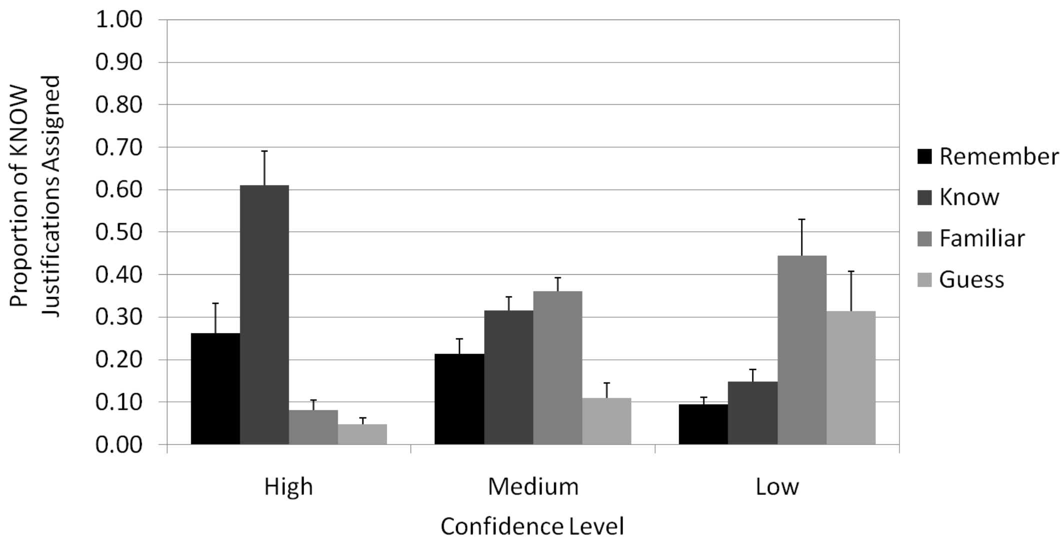

Figure 2.5. Proportion of KNOW justifications given High, Medium, and Low ratings of confidence in Experiment 2.1 that were assigned to Remember, Know, Familiar, and Guess. Error bars = 1 SeM.

For High confidence justifications, planned comparisons demonstrated that more KNOW justifications were assigned to Know than to Familiar, $t(3) = 6.22$, $p = .008$, or to Guess, $t(3) = 5.97$, $p = .009$; and more justifications were assigned to Remember than to Guess, $t(3) = 3.51$, $p = .039$. The difference between the proportion of High confidence KNOW justifications assigned to Know and Remember was not significant, $t(3) = 2.34$, $p = .10$, nor were the differences between the proportions assigned to Familiar and Remember, $t(3) = 2.26$, $p = .11$, or Familiar and Guess, $t(3) = 1.28$, $p = .29$.

For Medium confidence justifications, more KNOW justifications were assigned to Know than to Remember, $t(3) = 3.96$, $p = .029$, and to Familiar than to Guess, $t(3) = 9.55$, $p = .002$. The difference between the proportion of Medium confidence KNOW justifications assigned to Remember and Familiar was not significant, $t(3) = 2.18$, $p = .12$, nor were the differences between the proportions assigned to Familiar and Know, $t < 1$, or Remember and Guess, $t(3) = 1.59$, $p = .21$.

For Low confidence justifications, more were assigned to Familiar than to either Remember, $t(3) = 3.74$, $p = .033$, or Know, $t(3) = 3.41$, $p = .042$. The difference between the proportion assigned to Know and Remember approached significance, $t(3) = 2.99$, $p = .058$. The proportion of Low confidence KNOW justifications assigned to Guess was not significantly different from the proportion assigned to either Remember, $t(3) = 2.24$, $p = .11$, Know, $t(3) = 1.46$, $p = .24$, or Familiar, $t < 1$.

For these original KNOW justifications, ANOVAs comparing subjective experience responses across confidence levels demonstrated that a significant main effect of confidence level was evident for Know, $F(2,9) = 19.55$, $p = .001$, Familiar, $F(2,9) = 12.06$, $p = .003$, and Guess, $F(2,9) = 5.58$, $p = .027$, responses and Remember responses approached a significant main effect of confidence, $F(2,9) = 3.45$, $p = .077$. Planned comparisons demonstrated that the likelihood of a KNOW justification being assigned to Know fell with confidence level. A greater proportion of justifications were assigned to Know if they were High confidence compared to Medium confidence, $t(6) = 3.39$, $p = .015$, and if they were Medium confidence compared to Low confidence, $t(6) = 3.82$, $p = .009$. This pattern was reversed for the likelihood of a KNOW justification being assigned to Familiar. A greater proportion of Medium confidence justifications were assigned to Familiar than were High confidence justifications, $t(6) = 7.00$, $p < .001$, however the difference between the proportions of Medium and Low confidence KNOW justifications assigned to Familiar was not significant, $t < 1$. Finally, the likelihood of a KNOW justification being assigned to Remember was greater for Medium confidence KNOW justifications than for Low confidence KNOW justifications, $t(6) = 3.10$, $p = .02$. No other comparisons were significant.

These results demonstrate that for KNOW justifications the confidence levels associated with the justifications strongly influenced how participants assigned justifications to categories of subjective experience. As confidence level increased, the proportion of KNOW justifications assigned to Know increased, as did the proportion assigned to Remember. In addition, the proportion of KNOW justifications assigned to Familiar and Guess fell as confidence level increased.

2.3.3.4. Assignment of High, Medium, and Low confidence FAMILIAR justifications to Remember, Know, Familiar, and Guess

The pattern of categorisation of FAMILIAR justification statements parallels that for REMEMBER justifications shown in Figure 2.4. As is shown in Figure 2.6, between .50 and .60 of statements originally justifying a FAMILIAR recognition decision were appropriately assigned to the Familiar category of subjective experience irrespective of whether they were of High, Medium, or Low confidence.

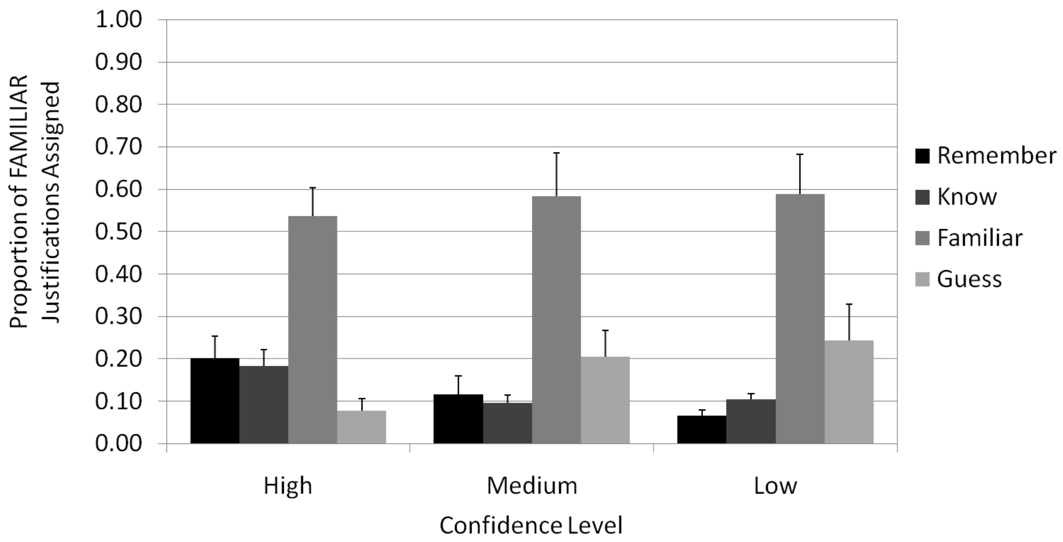

Figure 2.6. Proportion of FAMILIAR justifications given High, Medium, and Low ratings of confidence in Experiment 2.1 that were assigned to Remember, Know, Familiar, and Guess. Error bars = 1 SeM.

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

2.3.3.5. Assignment of High, Medium, and Low confidence GUESS justifications to Remember, Know, Familiar, and Guess

As is shown in Figure 2.7 the patterns for how GUESS justifications were assigned to subjective experience category again differed according to confidence level. Justifications given a Low confidence rating in Experiment 2.1 were appropriately assigned to Guess over .85 of the time, however when the confidence level was Medium, over .40 of justifications were assigned to both Familiar and Guess, and for High confidence GUESS justifications the majority (.48) were inappropriately assigned to Familiar. A 3(confidence level) x 4(subjective experience response) ANOVA on GUESS justifications demonstrated a significant main effect of subjective experience response, $F(1.38,12.42) = 20.30$, $p < .001$, and a significant interaction between subjective experience and confidence level, $F(1.38,12.42) = 12.45$, $p = .001$ (both Greenhouse-Geisser corrected). Separate ANOVAs at each level of confidence demonstrated significant main effects of subjective experience response at High, $F(3,9) = 6.49$, $p = .012$, and Low, $F(1.03,3.09) = 790.68$, $p < .001$ (Greenhouse-Geisser corrected), confidence levels, and the ANOVA for Medium confidence justifications approached significance, $F(3,9) = 3.53$, $p = .06$.

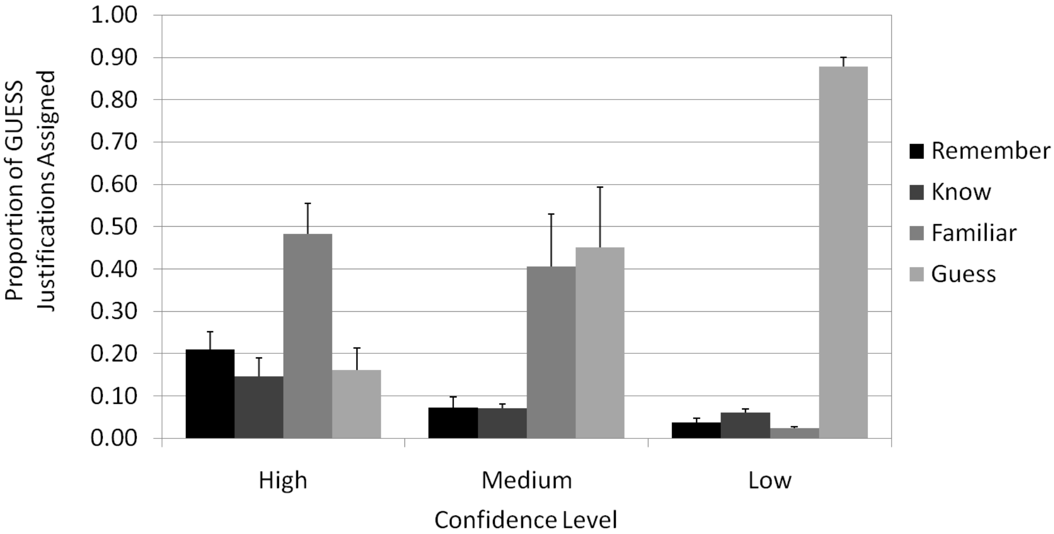

Figure 2.7. Proportion of GUESS justifications given High, Medium, and Low ratings of confidence in Experiment 2.1 that were assigned to Remember, Know, Familiar, and Guess. Error bars = 1 SeM.

Planned comparisons demonstrated that for Low confidence justifications more GUESS justifications were assigned to Guess than to either Remember, Know, or Familiar (all $p < .001$); in addition more justifications were assigned to Know than to either Remember, $t(3) = 10.11$, $p = .002$, or Familiar, $t(3) = 5.04$, $p = .015$, though the proportions here were very small. For Medium confidence GUESS justifications, results of planned comparisons did not reach conventional significance levels; more justifications were assigned to Familiar than to Remember, $t(3) = 2.96$, $p = .06$, or to Know, $t(3) = 2.80$, $p = .07$, and more justifications were assigned to Guess than to Remember, $t(3) = 2.37$, $p = .10$, or to Know, $t(3) = 2.58$, $p = .08$. For High confidence GUESS justifications, more justifications were assigned to Familiar than to Guess, $t(3) = 3.26$, $p = .047$, or Know, $t(3) = 3.00$, $p = .058$, or to Remember, $t(3) = 2.72$, $p = .073$, though these latter two comparisons did not reach significance. No other comparisons were significant.

ANOVAs comparing subjective experience response categories across confidence levels demonstrated that a significant main effect of confidence level was evident for Remember, $F(2,9) = 9.89$, $p = .005$, Familiar, $F(2,9) = 8.83$, $p = .008$, and Guess, $F(2,9) = 16.69$, $p = .001$ subjective experience responses and Know responses approached a significant main effect of confidence, $F(2,9) = 3.26$, $p = .086$. Examination of the patterns in Figure 2.7 suggested that the likelihood of a GUESS justification being assigned to Guess increased as confidence level fell. However planned comparisons demonstrated only the difference between proportion assigned to Low and Medium confidence levels was significant, with a greater proportion of GUESS justifications being assigned to Guess if they were Low confidence compared to Medium confidence, $t(6) = 2.97$, $p = .025$. This pattern was reversed for the likelihood of a GUESS justification being inappropriately assigned to Familiar. A greater proportion of Medium confidence GUESS justifications were assigned to Familiar than were Low confidence justifications, $t(6) = 3.09$, $p = .021$. In addition, the likelihood of a GUESS justification being inappropriately assigned to Remember was significantly higher if it was a High confidence justification as opposed to a Medium confidence justification, $t(6) = 2.79$, $p = .032$. No other comparisons were significant.

In sum, for GUESS justifications the confidence levels associated with the justifications strongly influenced how participants assigned these justifications to categories of subjective experience. As confidence level fell, the proportion of GUESS justifications assigned to Familiar became lower and the proportion assigned to Guess increased.

2.3.4. Discussion

This experiment aimed to explore lay-persons' understanding of the relationships between how one justifies a memory and the confidence associated with that justification. This was performed by examining how the confidence rating associated with a justification statement influenced whether that statement would be assigned to the appropriate category of subjective experience. For Remember or Familiar justification statements results demonstrated that, irrespective of these statements receiving a High, Medium, or Low, confidence rating by participants in Experiment 2.1, the majority of these justifications were appropriately assigned to Remember or Familiar categories of subjective experience respectively. For Remember justifications, an interaction between confidence level and subjective experience response was observed, with planned comparisons revealing that the likelihood of a Remember justification being inappropriately assigned to Familiar or Guess decreased as confidence level increased. For Familiar justifications, no interaction between confidence level and subjective experience response was observed. Conversely, the confidence levels assigned by participants in

Experiment 2.1 strongly influenced how Know and Guess justifications were assigned to categories of subjective experience. The proportion of Know justifications appropriately assigned to Know was found to increase as confidence level increased and conversely the proportion of Guess justifications appropriately assigned to Guess was found to increase as confidence level fell.

The suggestion that Remember justifications would be appropriately classified as Remember regardless of the confidence level of the justification was supported as at all levels of confidence the majority of justifications were assigned to Remember. Confidence level was only found to influence the assignment of a small proportion of Remember justifications to Familiar and Guess response categories. It is suggested that this is due to the nature of justifications made in support of recognition responses made on the basis of recollection. As Gardiner et al. (1998) reported, Remember justifications typically included details such as intra- and extra-list associations, item-related information, and self-reference. These rich, evocative details, which Gardiner et al. reported that participants were able to bring to mind with ease, are suggested as being critically important for identifying these items as Remember justifications. As is evident from examining the justifications employed in this experiment (shown in Appendix B.1) even the Remember justifications that had been given the lowest confidence ratings by participants in Experiment 2.1 included many such associations and details. For example, even the lowest confidence Remember justification involved intra-list association: *Ape "When it came up I remember thinking that there were lots of words with three letters".*

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

The justifications used in this experiment were those which had received the Highest, Medium, and Lowest confidence ratings for their subjective experience category in Experiment 2.1. However, examination of the Know and Familiar justifications used as stimuli in the current experiment (see Appendix B.1) highlighted differences between the types of reasons involved in justifications given High, Medium, and Low confidence ratings. High confidence Know justifications reflected certainty of recognition, e.g., *Bluebell "I am sure about that one, there were a couple of words which were similar and were part of the category flower"*; whereas Medium and Low confidence Know justifications mentioned absence of recollection, e.g., *Ring "I think I remember seeing it, but there was no link or image. I can't remember feeling anything"*; and one Low confidence justification mentioned familiarity. Some familiar justifications at each level of confidence mentioned familiarity (though not all justifications did), however only High and Low confidence Familiar justifications mentioned absence of recollection, and only Medium and Low justifications mentioned uncertainty, e.g., *Harbour "It was familiar, but I was confused. I knew it was there but could not be sure"*. As systematic differences across confidence levels were not evident to a greater extent for Know justifications than for Familiar justifications the content of the justifications cannot directly explain why confidence influenced classification of Know justifications but not Familiar justifications.

As with Know justifications, assignment of Guess justifications was also found to be dependent on confidence level. Examination of the High, Medium, and Low confidence Guess justifications selected for use in this experiment revealed that whereas Low confidence justifications were typically very short and explicitly discussed guessing, uncertainty or judgmental strategies, e.g., *Father "I kept saying "no" so I just guessed it was there because you said that 50% of the words were there"*; Medium and High confidence justifications also often mentioned confusion of certainty or familiarity of an item, e.g., *Flea "I am almost certain that it was there. But not entirely"*. It is suggested that these differences led participants in Experiment 2.1 to give these items different confidence ratings and participants in Experiment 2.2 to categorise a large proportion of High and Medium confidence Guess justifications inappropriately as Familiar. That some Guess responses reflect strategic decisions based on appropriate response rates while other guesses are based on lower levels of familiarity, or confusion of familiarity, appears to have here led participants to categorise these justifications differently.

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

Overall, these findings demonstrate that when understanding the memory experiences of others, participants’ judgments regarding subjective experience were influenced by confidence, and this influence was greater for Know and Guess categories of subjective experience. This novel method of using participants’ justification statements as stimuli has added to the experimental recognition findings of Rotello et al. (2005), McCabe and Geraci (2009), and Geraci et al. (2009), supporting the assertion that subjective experience judgments are influenced by manipulations of confidence. To further link this novel methodology to the prior experimental findings the confidence associated with each justification statement was manipulated more overtly in Experiment 2.3.

2.4. Experiment 2.3: Lay understanding of subjective experience when confidence is manipulated

2.4.1. Introduction

The aim of this experiment was to further explore the relationship between lay-persons' understanding of subjective experience and confidence by providing participants with both a justification statement and a confidence value on which to base their judgment of subjective experience. In this case, confidence was manipulated systematically in order to assess whether it influenced the subjective experience category the statement was assigned to. Prototypical Remember, Know, Familiar, and Guess justification statements were selected from around each category's mean confidence justification and were presented to participants accompanied by a confidence value that was either appropriate (to that subjective experience category) or inappropriate (from the confidence values of a different subjective experience category). As in Experiment 2.2 the task for participants was to assign justifications to the Remember, Know, Familiar, and Guess categories of subjective experience.

The focus of this experiment was whether participants based their categorisation decisions on both the justification and the confidence value, or whether one lent more to the categorisation decision than the other. If confidence value was ignored by participants this would result in uniform patterns of categorisation within subjective experience category; on the other hand, high or low confidence values may lead participants to interpret a statement as reflecting a different subjective experience to that which it was originally justifying. This experiment was interested in how people weigh the evidence provided in reports of experiential state and confidence ratings when interpreting the type of subjective experience being justified. In line with the findings of Experiment 2.2, it was predicted that Remember and Familiar justifications would be impervious to the confidence manipulation, but that appropriate classification of Know and Guess justifications would be influenced by the confidence level accompanying them.

2.4.2. Method

2.4.2.1. Participants

Data was collected using an online questionnaire active from October 2009 to February 2010. Advertisement was undertaken in the same manner as in Experiment 2.1 and Experiment 2.2. Full data sets from 258 participants were obtained (164 female, 92 male, 2 no response; mean age = 31.04, SD = 17.17, range = 16 to 79).

2.4.2.2. Materials and Design

The cue words and justification statements to be used as items in the online questionnaire were a selection from those published by Gardiner et al. (1998) and used in Experiment 2.1. Using the confidence ratings obtained from participants in Experiment 2.1, within each subjective experience category (Remember, Know, Familiar, and Guess) eight statements from around the mean confidence value were selected as prototypical justifications for each category (shown in Appendix B.1). Where cue words were duplicates, one of the cue words was exchanged for a word from the MRC Psycholinguistic database which matched the original on familiarity, no. of letters, and first letter, e.g., *bluebell* swapped for *blessing.*

In this experiment, as justifications were to be paired with confidence values, eight confidence values were selected around the mean for each subjective experience category and these are shown in Table 2.6. It was ensured that the ranges of values for each category were non-overlapping and that the mean of the selected values matched the original mean confidence as closely as possible[8].

In the pairing of items to confidence values, a Latin-square design was used. Within each of the four categories of subjective experience, two justifications were paired with plausible confidence values and two justifications were paired with confidence values derived from the confidence ranges for each of the other original subjective experience categories. For example, of the eight Remember justifications two were paired with Very High confidence values, two were paired with High confidence values, two with Medium confidence values,

[8] Means shown in Table 2.6 do not exactly match those in Table 2.4 as selection of items for this experiment was performed prior to final analysis of Experiment 2.1. Differences in means varies from 0.30% to 1.30%.

and two with Low confidence values. This pairing was then repeated for Know, Familiar, and Guess justifications. Using this design each participant saw 32 items. Four versions of the pairings were created. Careful pairing was done to ensure that each confidence value served as a plausible confidence once, for example a Very High confidence of 85 matched to a Remember justification, but on the other versions of the questionnaire that confidence value of 85 was paired with a either a Know, Familiar, or Guess justification. Each confidence value was only used once within each version of the questionnaire. Participants were randomly assigned to one of the four versions when they accessed the web page and within each version items were presented in random order.

Table 2.6. Means and ranges of confidence values for the justifications from each subjective experience category.

Subjective experience category	Mean confidence rating for category (from Exp. 2.1)	Range of confidence values provided to participants	Mean of range	Confidence level*
Remember	81.32	78 – 85	81.50	Very High
Know	56.88	53 – 60	56.50	High
Familiar	45.80	42 – 49	45.50	Medium
Guess	34.84	31 – 38	34.50	Low

* ‘Confidence level’ denotes the title used to refer to the confidence range throughout the following experimental report.

2.4.2.3. Procedure

The procedure for this experiment closely followed that of Experiment 2.2 though here participants were presented with both a justification statement and a confidence value on which to base their classification of the item as Remember, Know, Familiar, or Guess. The instructions given at the start of the online questionnaire regarding how the justifications statements were made by previous participants matched those given for Experiment 2.2. Current participants were then instructed that in this experiment they would be shown justification statements and confidence ratings made by previous participants in a memory test and that their task was to make a judgment about the person’s recognition decision by classifying it as Remember, Know, Familiar, or Guess. Full definitions of the categories were provided (see Appendix A.1) and further instructions matched those in Experiment 2.2 except participants were also reminded to pay attention to both the justification and the confidence rating for each item. Participants then entered their demographic information before being shown the first item. For each item, participants were shown the cue word, justification statement, and confidence rating, accompanied by the question “Which recognition category?” and the options Remember, Know, Familiar, and Guess. They made their decisions by clicking the appropriate button. After seeing all 32 items participants were debriefed as in Experiment 2.2.

2.4.3. Results

This experiment was interested in whether participants were able to reliably match justification statements to their original category of subjective experience when provided with a confidence value alongside the justification. Analysis in this experiment took the form of items-analysis with the original subjective experience of the item (REMEMBER, KNOW, FAMILIAR, GUESS) and the manipulated confidence level of the item (Very High, High, Medium, Low) being treated as between-items factors and the response of the participant – which category of subjective experience they assigned the justification to (Remember, Know, Familiar, or Guess), being treated as a within-items factor. The dependent variable was the proportion of items within an original subjective experience category and a particular level of confidence that were assigned to each of the subjective experience response categories. For example, the proportion of REMEMBER justifications accompanied by a Very High level of confidence that were assigned to Remember, Know, Familiar, and Guess. Proportions were calculated across participants and firstly a 4(original subjective experience) x 4(manipulated confidence level) x 4(subjective experience response) ANOVA was performed. Separate ANOVAs and further comparisons were then performed to examine patterns within the different levels of confidence and original subjective experience categories.

The 4(original subjective experience) x 4(manipulated confidence level) x 4(subjective experience response) ANOVA demonstrated a significant main effect of subjective experience response, $F(2.02,226.42) = 35.22$, $p < .001$, and a significant interaction between subjective experience response and original subjective experience, $F(6.09,226.42) = 102.45$, $p < .001$ (both Greenhouse-Geisser corrected). This interaction suggests that the pattern of responses across subjective experience categories differed depending on the original subjective experience of the item. For this significant interaction, separate ANOVAs and further comparisons were conducted to examine patterns within the original subjective experience categories and different subjective experience responses; these are reported in Section 2.4.3.2. The interaction between subjective experience response and manipulated confidence level was not significant, $F < 1$, and neither was the three-way interaction between original subjective experience, manipulated confidence level, and subjective experience response, $F < 1$. The non-significant three-way interaction is discussed first, followed by the two-way interactions.

2.4.3.1. Assignment of justifications to Remember, Know, Familiar, and Guess when confidence was manipulated

As was demonstrated by the lack of an interaction between manipulated confidence level and subjective experience response or a three-way interaction between these two variables and original subjective experience, in general, manipulated confidence level did not influence how participants assigned the REMEMBER, KNOW, FAMILIAR, and GUESS justifications. This is illustrated by Figure 2.8, Figure 2.9, Figure 2.10, and Figure 2.11, which show for each of the four original subjective experience categories the proportion of justifications that were categorised as Remember, Know, Familiar, and Guess split by manipulated confidence level.

Figure 2.8 shows that for statements originally justifying a REMEMBER recognition decision, over .50 were appropriately assigned to the Remember category of subjective experience regardless of what level of confidence the statement was paired with on the questionnaire. Approximately .25 were assigned to Know, .10 to Familiar, and under .05 to Guess. For KNOW justification statements, Figure 2.9 shows that regardless of what level of confidence the statement was paired with on the questionnaire approximately equal proportions of justifications (between .28 and .38) were assigned to Know and Familiar. Likewise for FAMILIAR justification statements, Figure 2.10 demonstrates that at all levels of manipulated confidence approximately .50 of justifications were appropriately assigned to the Familiar category of subjective experience, with approximately .25 assigned to Guess. Finally, for statements originally justifying a GUESS recognition decision, Figure 2.11 shows that at all levels of manipulated confidence around .40 of justifications were assigned to both Familiar and Guess.

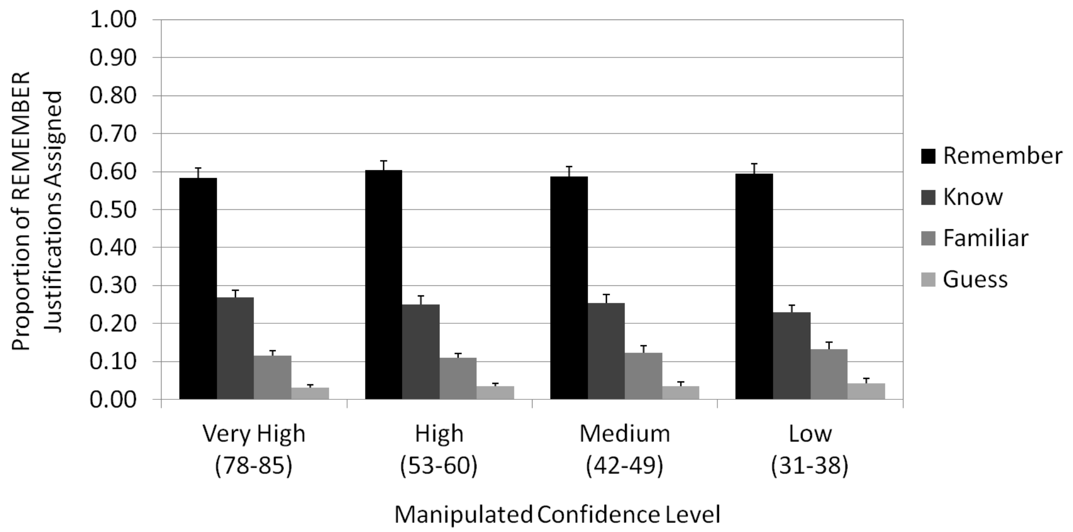

Figure 2.8. Proportion of REMEMBER justifications matched with Very High, High, Medium, and Low confidence values that were assigned to Remember, Know, Familiar, and Guess. Error bars = 1 SeM.

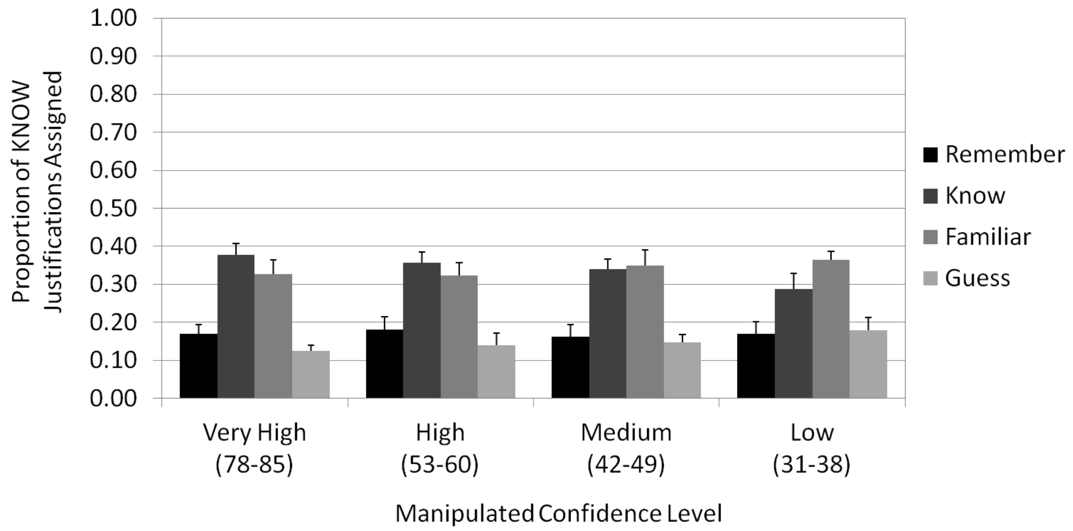

Figure 2.9. Proportion of KNOW justifications matched with Very High, High, Medium, and Low confidence values that were assigned to Remember, Know, Familiar, and Guess. Error bars = 1 SeM.

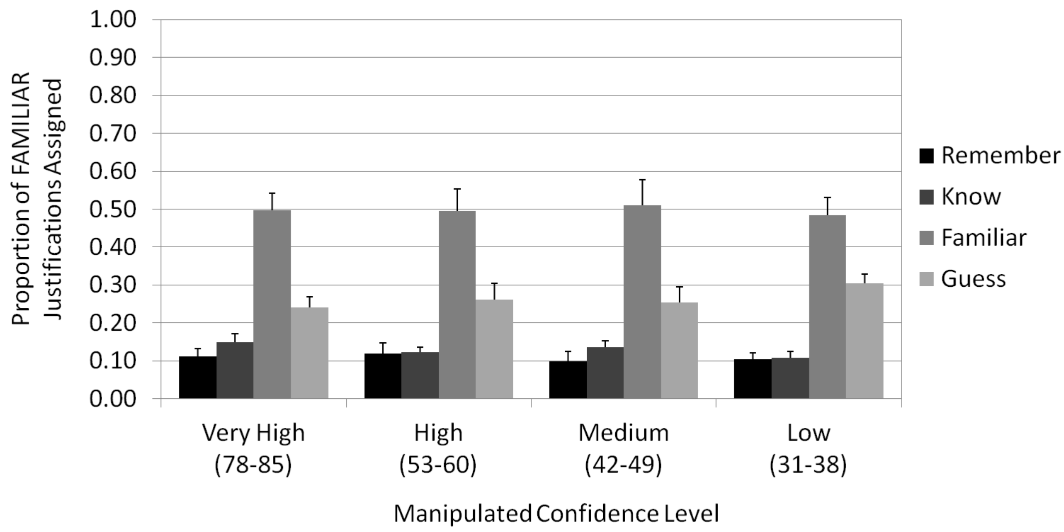

Figure 2.10. Proportion of FAMILIAR justifications matched with Very High, High, Medium, and Low confidence values that were assigned to Remember, Know, Familiar, and Guess. Error bars = 1 SeM.

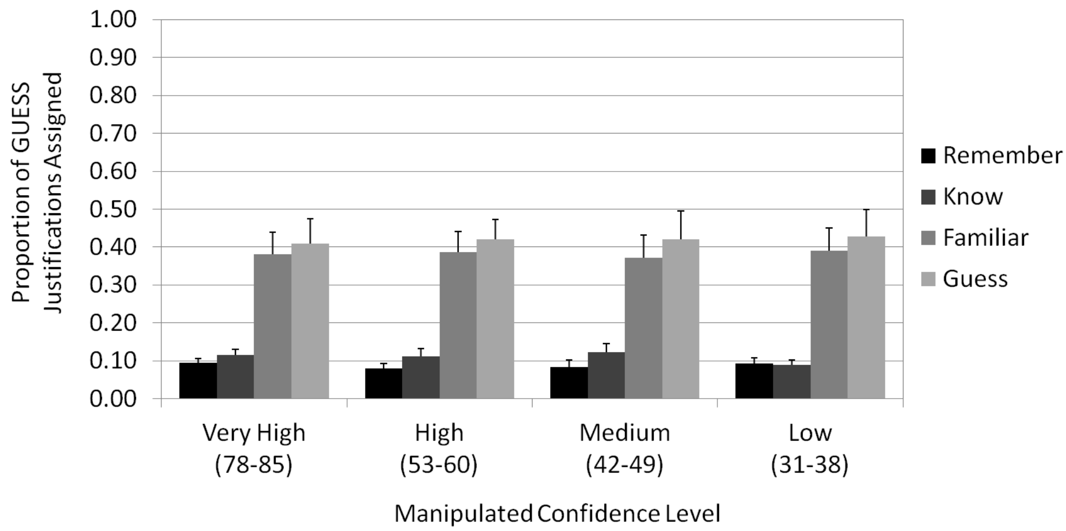

Figure 2.11. Proportion of GUESS justifications matched with Very High, High, Medium, and Low confidence values that were assigned to Remember, Know, Familiar, and Guess. Error bars = 1 SeM.

In sum, these results demonstrate that patterns of responding were stable across confidence level. The confidence value that was provided alongside the justification statement did not influence how participants assigned the item to a category of subjective experience; participants primarily based their assignment of the item to Remember, Know, Familiar, or Guess on the experiential state reported in the justification statement.

2.4.3.2. Analysis of two-way interactions

As discussed above, there was no significant interaction between manipulated confidence and subjective experience response. This is demonstrated by the means in Figure 2.12, which show that, regardless of what the original subjective experience category the statement was justifying, the proportion of justification statements assigned to each of the subjective experience response categories were very similar across the levels of manipulated confidence. Over .30 were consistently assigned to Familiar, around .22 to Remember, and around .20 to both Know and Guess.

In contrast, there was a significant two-way interaction between subjective experience response and original subjective experience category. Means for this interaction are shown in Figure 2.13. This figure shows very similar patterns to those shown for this interaction in Experiment 2.2, shown in Figure 2.3. When manipulated confidence level is not considered, the majority of REMEMBER and FAMILIAR justification statements were appropriately allocated to their original subjective experience category. For KNOW justifications nearly as many were assigned to Familiar as to Know, and for GUESS statements nearly as many were assigned to Familiar as to Guess.

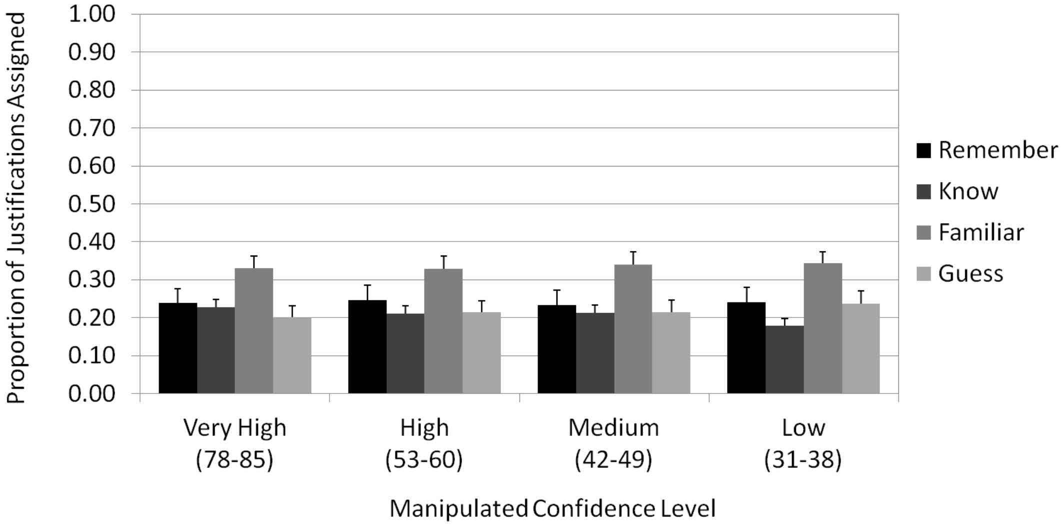

Figure 2.12. Proportion of justifications at each level of manipulated confidence that were categorised as Remember, Know, Familiar, and Guess irrespective of original subjective experience category. Errors bars = 1 SeM.

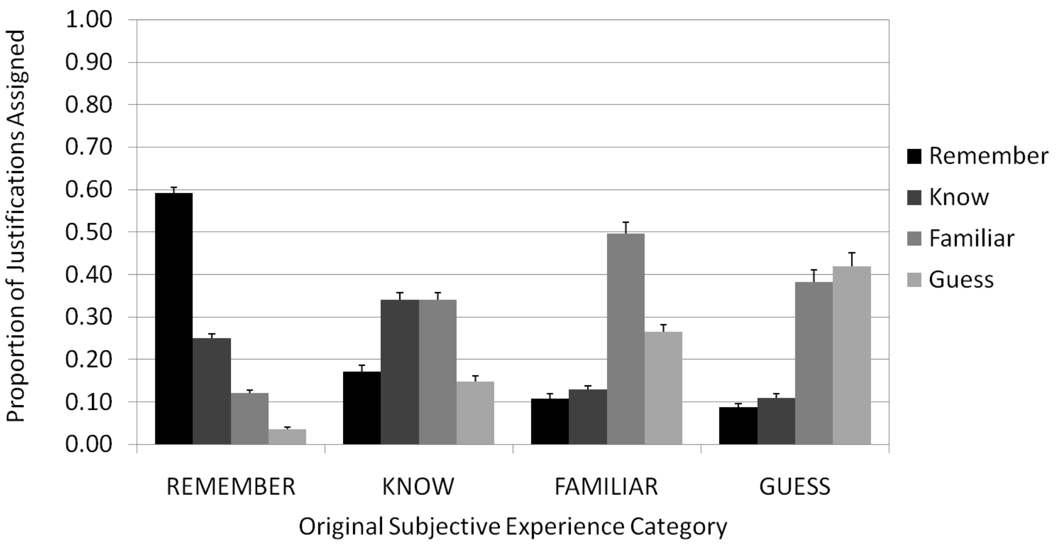

Figure 2.13. Proportion of justifications of the four original subjective experience categories that were categorised as Remember, Know, Familiar, and Guess regardless of manipulated confidence level. Errors bars = 1 SeM.

Further comparisons were performed for this significant two-way interaction. Within each original subjective experience category ANOVAs demonstrated significant within-subjects effects of subjective experience response. In each original subjective experience category, proportion of justifications assigned to the four subjective experience responses differed: REMEMBER, $F(1.97,61.16) = 546.10$, $p < .001$, KNOW, $F(2.34,72.49) = 38.85$, $p < .001$, FAMILIAR, $F(1.41,43.64) = 81.48$, $p < .001$, and GUESS, $F(1.24,38.50) = 48.44$, $p < .001$. Planned comparisons were carried out between each of the subjective experience responses within the four original subjective experience categories.

For REMEMBER justifications, all comparisons were significantly different (all $p < .001$); a significantly greater proportion of justifications were assigned to Remember compared to all other categories, Know compared to Familiar and Guess, and Familiar compared to Guess. For KNOW justifications, comparisons between the proportion of justifications assigned to Know and Familiar, and Remember and Guess were not significant (both $t < 1$); .34 of KNOW justifications were assigned to both Know and Familiar and .17 to Remember and .15 to Guess. All other comparisons for KNOW justifications were significant (all $p <$ .001). For FAMILIAR justifications, the comparison between the proportion assigned to Remember and Know was not significant, $t(31) = 1.54$, $p = .13$; all other comparisons were significant (all $p < .001$), a significantly larger proportion of FAMILIAR justifications were appropriately assigned to Familiar than to any other response category. Finally, for GUESS justifications, no significant difference was found between the proportion of justifications assigned to Familiar and Guess, $t < 1$; .38 of GUESS justifications were assigned to Familiar and .42 to Guess. All other comparisons for GUESS justifications were significant (all at least $p < .02$); significantly more GUESS justifications were assigned to Familiar and Guess than to either Remember and Know, and more were assigned to Know than to Remember, though the proportion assigned to each of these was below .11.

These results demonstrate that when confidence is not considered, REMEMBER and FAMILIAR justification statements were appropriately allocated to their original subjective experience category whereas assignment of KNOW and GUESS justifications to their respective categories was not as clear-cut; nearly as many KNOW and GUESS justifications were assigned to Familiar as were appropriately categorised.

ANOVAs across the four original subjective experience categories demonstrated significant between-items effects for the four subjective experience response categories. For each response category, the proportion of justifications assigned to that category differed according to the original subjective experience category of the item: Remember, $F(3,124) = 414.65$, $p < .001$, Know, $F(3,124) = 89.59$, $p < .001$, Familiar, $F(3,124) = 56.75$, $p < .001$, and Guess, $F(3,124) = 70.82$, $p < .001$. Planned comparisons were carried out for each of the subjective experience responses across the four original subjective experience categories. Comparing across REMEMBER and KNOW original subjective experience categories of justifications, a larger proportion of REMEMBER justifications were assigned to a Remember response than were KNOW justifications; and conversely a higher proportion of KNOW justifications were assigned to Know, Familiar, and Guess than were REMEMBER justifications (all $p < .001$).

Comparing across REMEMBER and FAMILIAR original subjective experience categories, a larger proportion of REMEMBER justifications were assigned to Remember and Know than were FAMILIAR justifications; conversely a higher proportion of FAMILIAR justifications were assigned to Familiar and Guess than were REMEMBER justifications (all $p < .001$). This pattern was repeated for the comparison of REMEMBER and GUESS original categories of justifications (all $p < .001$).

Comparing across KNOW and FAMILIAR original subjective experience categories, a larger proportion of KNOW justifications were assigned to Remember and Know than were

FAMILIAR justifications; conversely a higher proportion of FAMILIAR justifications were assigned to Familiar and Guess than were KNOW justifications (all at least p = .001). This pattern was repeated for the comparison of KNOW and GUESS original categories of justifications, though here there was no difference between the proportion of KNOW and GUESS justifications that were inappropriately assigned to Familiar (.34 and .38 respectively), $t(62) = 1.31$, $p = .20$. All other comparisons $p < .001$.

Comparing across FAMILIAR and GUESS original subjective experience categories, a larger proportion of FAMILIAR justifications were assigned to Familiar than were GUESS justifications; conversely a higher proportion of GUESS justifications were appropriately assigned to Guess than were FAMILIAR justifications (both at least p = .004). There were no significant differences between the proportion of FAMILIAR and GUESS justifications that were inappropriately assigned to Remember, $t(62) = 1.52$, $p = .13$, or to Know, $t(62) = 1.58$, $p = .12$.

In sum, these results demonstrate that reliable patterns of data are observed both within and across the original subjective experience categories the justifications belonged to. The proportion of justifications categorised as Remember, Know, Familiar, or Guess were highest when the justification was originally a REMEMBER, KNOW, FAMILIAR, or GUESS justification respectively, i.e., of the justifications assigned to each response option the majority come from the appropriate original subjective experience category. Whilst some KNOW and GUESS justifications were allocated to Familiar instead of to their appropriate category, the largest proportion of justifications that were assigned to Know and Guess were KNOW and GUESS justifications.

2.4.3.3. Anecdotal reports

After completing the questionnaire participants provided comments on what they thought of the questionnaire. Some comments demonstrated that, in line with the findings from the experiment as a whole, the participant had used the justifications to a much greater extent than the confidence values when assigning the item to Remember, Know, Familiar, or Guess. In addition to the comment provided as a quote at the start of this chapter, three further example comments are:

> *The most important criterion is the description of the mental process behind the recognition. If it is associated with something else in the surrounding then*

is 'Remembering'. If there is only the knowledge then is 'Know'. If it is neither but the person is more or less sure to having seen it (and the score must reflect that) then it is 'Familiar'. It is 'Guess' when no other case feels right and/or the score is very low.

I tried to make my category decisions on how certain the participant sounded in their verbal answer. Vague answers like, it feels like it was there, I ranked as 'familiar', and more definite descriptions of how they remembered the word I ranked as 'know'. I found some of the confidence scores conflicted with my interpretation of the participants statement, but I found the statements more compelling.

I thought that the certainty ratings were very subjective, and depended too much on the participants' feeling at the time (some people who basically justified their response as guesses seemed to have ratings that were higher than ratings where people were fairly certain they remembered). So, my decisions took into account the ratings, but put more weighting on justifications.

This anecdotal evidence supports the experimental data and highlights that some participants were also able to reflect on the process of using both a justification statement and a confidence value to categorise an item to a subjective experience category.

2.4.4. Discussion

In this experiment, participants were provided with a confidence value alongside the justification statement that they were to assign to a category of subjective experience. Results demonstrated that participants' assignment of justifications to subjective experience categories was not influenced by confidence. There was no three-way interaction between original subjective experience, manipulated confidence level, and subjective experience response, and no two-way interaction between subjective experience response and manipulated confidence level. Participants focused their categorisation on the experiential reports provided by the justification statements; they did not seem to use the provided confidence value when categorising the justifications as

Remember, Know, Familiar, or Guess. Furthermore, the anecdotal reports provided by some participants at the end of the questionnaire revealed that they had insight into their use of the justification statement in preference to the confidence value as a basis for their categorisation.

With regard to the patterns of responses, as in Experiment 2.2 Remember and Familiar justifications were consistently appropriately categorised as Remember or Familiar whilst Know and Guess statements were often inappropriately categorised as Familiar. Examination of the Know justifications utilised in this experiment revealed that they typically referred to lack of recollection or uncertainty about recognition of the word; for example, *Tangerine "I recognised it as a word from yesterday, but I cannot really remember what I thought, I could not remember seeing it on the screen but I was sure it was there yesterday".* None of the Know statements included the word 'familiarity' so it was not the use of this word which led to Know justifications being inappropriately assigned to Familiar. Instead it could be suggested that the uncertainty evident in some of the Know justifications (see *Tangerine* above) is what led participants to categorise some justifications as Familiar instead of Know. This could suggest that the separation of traditional Know category into separate Know and Familiar categories could require more refinement of definitions; however, evidence from previous research that has separated these categories (e.g., Brewer & Sampaio, 2006; Conway et al., 1997; Dewhurst et al., 2009; Herbert & Burt, 2001, 2003, 2004; Kihlstrom et al., 1996) and other experiments presented in this thesis suggests that many aspects of this split are reliable. This issue will be returned to in the General Discussion (Section 2.6.1).

Examination of the Guess justifications revealed that all statements utilised in this experiment referred to uncertainty about recognition; for example, *Harbour "I lived by the sea all my life, so I was not sure whether I have encountered that word here or whether it is to do with home".* Only one of the eight Guess justifications included the word 'familiar'. Furthermore, while Gardiner et al. (1998) demonstrated that Guess justifications can often reflect inferences or judgmental strategies (see Section 2.1.2), only one of the statements used in the current experiment contained an inference concerning types of words from the study list: *Harp "It seemed that there were quite a few musical instruments, so I took a guess that it came up".* Instead of judgmental strategies and inferences, it is suggested that the uncertainty about why an item felt familiar inherent in the Guess justifications used in this experiment is what led many of these justifications to be categorised as Familiar.

In conjunction with the findings from Experiment 2.2, the current findings demonstrate that when confidence is manipulated as an internal attribute of the justification statement (Experiment 2.2) it influences how the justification is assigned to subjective experience category, especially for Know and Guess justifications. However, when confidence is manipulated as a linked but external feature of the justification (Experiment 2.3) it does not influence how justifications are assigned. Tentatively this could be taken as support for the view that confidence derives from subjective experience as proposed by Tulving (1985) and Gardiner "…it is surely the subjective state of awareness that gives rise to confidence in memory, not confidence that gives rise to the state of awareness" (Gardiner, 2001, p. 1356). When both a statement and a confidence value were provided, categorisation appeared to be largely based on the justification report rather than the confidence value. It could be suggested that this occurred because participants used their lay understanding of memory experiences to determine that subjective experience is of more importance than confidence.

An alternative suggestion could be that participants overlooked the confidence value in making their subjective experience judgment as they guessed the purpose of the experiment, or thought the two items were not linked, or the justification included more details and therefore occupied their attention. To follow this up, future experiments could manipulate confidence by providing a confidence value but could make it internal to the justification statement, for example *"I was about 30% confident because I remembered there had been animals on the list of words but I wasn't sure if elephant had been there"*, or *"I knew I'd thought about a 'frog' yesterday so I thought it was there, I'd say I was about 42% confident".* In the current experiments it was considered important to remain true to Gardiner et al.'s (1998) justification statements as they were real justifications obtained from actual participants. However, in future, if justifications were carefully worded and thorough matching was performed, manipulating confidence within the justification statement would be a valuable line of enquiry.

In sum, the current findings demonstrate that when a confidence value is provided alongside a justification statement this does not influence lay persons' understanding of memory experiences. Remember justifications were classified as Remember, Familiar justifications were classified as Familiar, Know justifications were classified as either Know or Familiar, and Guess justifications were classified as either Guess or Familiar. The different patterns of categorisation of Know and Familiar justifications demonstrate that there is something within the justifications which enables people to categorise them differently. The memory processes that are reflected in these justification types were the focus of investigation in Experiment 2.4.

2.5. Experiment 2.4: Lay-person categorisation of Know justifications into Know and Familiar

2.5.1. Introduction

The previous experiments have shown that Know and Familiar justification statements receive different ratings of confidence (Experiment 2.1), demonstrate different patterns of assignment to categories of subjective experience (Experiments 2.2 and 2.3), and their patterns of assignment are differentially influenced by the confidence associated with the justification (Experiment 2.2). The aim of the current experiment was to test whether non-experts could observe the differences between statements justifying a Know subjective experience and a Familiar subjective experience as a test of the validity of the Know-Familiar split. Crucially, here participants were not asked to match the justification statement to its category based on experimental definitions, participants were simply asked to split the 90 original Gardiner et al. (1998) Know justifications into two different types. It was predicted that, based purely on intrinsic understanding of memory experiences, participants would recognize two different types of subjective experience from the justifications made by others to items recognised on a memory test. If people have access to information at retrieval which maps onto the states of Knowing and Familiarity, *and* which is reflected in their statements made when asked to justify their recognition, it might be expected that other people are able to recognise those mnemonic processes and identify that different justifications are a sign of different underlying experiential states.

2.5.2. Method

2.5.2.1. Participants

Participants were 50 psychology students (41 female) from the University of Leeds. Individual ages were not recorded but participants confirmed that they were aged between 18 and 35 on the consent form. Undergraduate students received Participant Pool Credits for taking part. Participants were tested in groups of between two and ten in a classroom. One participant failed to follow experimental instructions and her data were excluded from analysis.

2.5.2.2. Materials and Design

The cue words and justification statements to be used as items in the online questionnaire were those 90 classified as Know justifications by participants in the experiment by Gardiner et al. (1998) which had previously been expertly categorised as Familiar (n = 43) and Know (n = 47; see Section 2.1.4). Each cue word and its justification statement were printed on small cards and each participant was given an envelope with the 90 cards in at the start of the experiment. A debrief questionnaire asking participants what criteria they had based their sorting decisions on was also created for participants to complete at the end of the experiment (see Appendix F).

2.5.2.3. Procedure

Participants were given the experiment instructions along with an envelope and a set of justification cards. They were informed that no further instructions would be provided so they should read the instruction sheet carefully. The instructions specified that this experiment was interested in whether participants could differentiate between two types of memory justification statement. Participants were told that in the envelope were 90 cards on which were written 90 cue words and statements and that their task was to sort these into Type A and Type B statements. The statements were defined as justifications of memory provided by participants in a previous memory experiment with the following explanation: “on a previous memory test participants had said yes, they recognised that cue word, and the justification on the card is the reason they gave for *why* they thought they recognised that cue word.”

Participants were instructed to read the statements and then begin sorting them into Type A and Type B statements along whatever criteria they thought the statements differed. The instructions specified that there were no right or wrong answers, and participants did not have to end up with equal numbers of Type A and Type B statements. Participants were instructed to place Type A cards in one pile and Type B cards in another pile, check that they were happy with their sorting decisions as many times as needed, and write their initials and 'Type A' or 'Type B' on the back of the cards before placing the cards back in the envelope and completing the decision criteria questionnaire. In this questionnaire participants were asked how they made their decisions, how easy it was to sort the cards, how similar or different they thought their Type A and Type B statements were, and how confident they thought the participants who had made the Type A and B statements had been. Finally participants were shown definitions of Know and Familiar and asked how their Type A and B statements mapped onto those definitions (see Appendix F for full questionnaire).

2.5.3. Results

Analysis of the data from this experiment was conducted in a number of different ways. Firstly, to test the association between participant and expert classification of Know and Familiar justifications, Goodman-Kruskal's gamma correlations (Goodman & Kruskal, 1954) were used. Gamma is a non-parametric measure of association which, in the current experiment, was used to assess the likelihood that a justification labelled as Know by experts would be labelled as Know by participants and not as Familiar (see Appendix E for full explanation of the gamma correlation formula). Subsequent analysis explored the proportion of items that had been appropriately classified as Know and Familiar and participants' responses on the decision criteria questionnaire were analysed to see on what basis participants' reported sorting the justifications.

2.5.3.1. Mapping of Know and Familiar to Type A and Type B labels.

In this experiment, instead of participants being provided with definitions of Know and Familiar subjective experiences and asking them to match the memory justification statements to these labels they had simply been asked to sort the justifications along any criteria on which they thought they differed and then label them as Type A and Type B. Consequently, it was not known whether their conception of what defined a Type A justification mapped onto our definition of Know or whether they had labelled Know justifications as Type B, and so on. To ensure that the mapping of participant Type A and B labels to the expert Know-Familiar categorisation was accurate it was conducted in three ways. Each method resulted in four values for each participant – the number of Know justifications that were labelled Type A, the number of Know justifications that were labelled Type B, the number of Familiar justifications that were labelled Type A, and the number of Familiar justifications that were labelled Type B.

Firstly, at the end of the decision criteria questionnaire participants were provided with definitions of Know and Familiar and were asked which fit their Type A and B justifications. However, eight participants did not respond appropriately to this question: four did not respond at all and four put the same response for both Know and Familiar definitions, i.e., reported that the definition of Know fit their Type A *and* their Type B statements. For the remaining 41 participants the first method used for mapping of Type A and B to Know and Familiar was to use their questionnaire responses. For each participant the number of Know and Familiar justifications "correctly" and "incorrectly" categorised were then calculated.

Using this first method of mapping Type A and B to Know and Familiar it became apparent that, for some participants, although their questionnaire responses stated that, for example, they thought their Type A statements mapped onto the Know definition and their Type B to Familiar, their actual sorting results demonstrated that the majority of their Type A pile were Familiar justifications, suggesting that their conceptualisation of Type A actually mapped onto the Familiar definition. This is reasonable as participants were only asked to label their piles as A and B after they had finished sorting the statements into two types, and were only asked to link their A and B to the definitions of Know and Familiar at the end of the decision criteria questionnaire.

The second method used for mapping of Type A and B to Know and Familiar was to base the classification around the majority response. For example, looking at the data for Participant A in Table 2.7, as the largest value of this participants four response options is 44 for number of Know justifications labelled as Type A, for this participant their label Type A would be considered to map onto Know and accordingly their label Type B would be considered to map onto Familiar. Not all participants' data was this 'clean' however; for example, mapping of Type A to Know and Type B to Familiar would be the same for Participant B in Table 2.7. Even though for this participant the majority of both the Type A and Type B justifications were actually Know justifications this method of mapping was based around the single largest value which here is 27 Know justifications labelled as Type A.

Table 2.7. Example data of Know and Familiar justifications sorted into Type A and Type B.

	Participant A's Response		Participant B's Response	
Correct Answer	Type A	Type B	Type A	Type B
Know	44	3	27	20
Familiar	13	30	25	18

For 66% of the 41 participants who provided appropriate questionnaire responses, how Type A and B mapped onto Know and Familiar based on their questionnaire response matched the mapping based on the majority response. However, there were 14 participants for whom their questionnaire mapping of Type A and B to Know and Familiar did not match the mapping based on the majority response, data for these participants was analysed separately.

2.5.3.2. Association between participant and expert categorisation of Know and Familiar justifications

Using each of the above methods of mapping Type A and Type B responses onto Know and Familiar definitions, Goodman-Kruskal (1954) gamma correlations were calculated for each participant. This correlation results in a score of between -1 and +1. The more positive this correlation coefficient the stronger the association between participant and expert classification of Know and Familiar justifications. Mean gammas were calculated across participants and are shown in Table 2.8. Each of these gammas was subjected to a one-sample *t*-test and all were found to be significantly different from 0 (all at least $p < .004$). Comparing the mean gammas in the top two rows of Table 2.8 it is evident that when calculating the association between participant and expert classification of Know and Familiar justifications, if this was based on participants' questionnaire response the resulting gamma was lower (.26) than when calculated on majority response (.54) and this difference was found to be significant, $t(40) = 3.51$, $p = .001$. When calculated using majority response the association between participant and expert classification of Know and Familiar justifications was significantly stronger than when calculation was based on questionnaire response. This suggests participants were good at splitting the justifications into their respective Know and Familiar categories but some were not so good at articulating their reasoning on the decision criteria questionnaire.

As is shown in the bottom two rows of Table 2.8 however, by splitting this analysis by whether the participants' questionnaire response had matched their majority response it becomes clear that the lower overall gamma based on questionnaire response was due to a negative mean gamma (-.41) being obtained for the 14 participants whose questionnaire response did not match the response based on the majority value.

Table 2.8. Mean gamma correlation coefficients (and standard deviations) calculated using different methods of mapping Type A and B to Know and Familiar definitions.

Mapping of A and B to Know and Familiar based on	N	Gamma
Questionnaire response (all valid responses)	41	.26 (.55)
Majority response	49	.54 (.28)
Questionnaire response (where questionnaire response matched majority response)	27	.61 (.25)
Questionnaire response (where questionnaire response did not match majority response)	14	-.41 (.29)

For these participants, how they had mapped their Type A and B statements onto Know and Familiar (based on their questionnaire response) lay in opposition to the expert classifications of Know and Familiar. There was no objective truth to their mapping of A and B to Know and Familiar. Because of this problem with some participants providing, what were deemed to be, incorrect questionnaire responses regarding how their Type A and Type B statements mapped onto the Know and Familiar categories, all subsequent analysis was conducted using mapping based on the majority response.

2.5.3.3. Proportion categorised correctly

The second method of analysis conducted on this data was calculation of the proportion of Know and Familiar justifications that were "correctly" labelled as Know and Familiar. For Know justifications the mean proportion categorised as Know was .74 (SD = .17), which was significantly higher than the mean proportion of Familiar justifications that were assigned to Familiar, M = .52 (SD = .21), $t(48) = 4.35$, $p < .001$. Participants were significantly better at appropriately categorising Know justifications than Familiar justifications. One-sample t-tests comparing these mean performance scores to chance

(.50) revealed that performance for categorisation of Know justifications was significantly above chance, $t(48) = 10.02$, $p < .001$, however performance for categorisation of Familiar justifications was not, $t(48) = .68$, $p = .50$. In sum, participants were reliably able to assign Know justifications correctly to Know but were not consistently able to correctly categorise Familiar justifications[9].

2.5.3.4. Decision criteria

After sorting of the statements into Type A and Type B was completed, participants were instructed to fill in a questionnaire asking about the criteria they used to sort the justifications. Firstly, participants were asked to rate how easy it was to sort the statements, how similar or dissimilar they found the statements, and whether they thought other people would sort the statements in the same way as them. For each question a five-point scale was used from -2 to +2. Mean scores are shown in Table 2.9. One-sample t-tests against the mid-point score of 0 demonstrated that participants rated the statements as significantly difficult to sort, $t(48) = 5.54$, $p < .001$; thought Type A and B statements were significantly similar as opposed to different, $t(48) = 3.50$, $p = .001$; and considered it neither unlikely or likely that others would sort the statements in the same way as them, $t < 1$.

Table 2.9. Mean scores and standard deviations for ratings of ease, similarity, and other peoples' sorting.

Question	Mean score	Standard deviation
Ease of sorting	-.67	.85
Similarity of statements	.47	.94
Other people make same sorting decisions	.02	.95

Participants were then shown a list of possible criteria on which they may have made their sorting decisions and were asked to tick those which they had used. The majority of reasons were phrased as a comparison of statements, e.g., for 'More sure': *I thought that for one Type of memory statement the people making the statements sounded more sure of their memories than for the other Type of memory statement.* The full questionnaire can be seen in Appendix F. The percentages of participants who endorsed each of the reasons

[9] This analysis was repeated in the form of items-analysis and is shown in Appendix G.

are shown in Table 2.10. The most common criterion endorsed was the above example relating to how sure the people making the justifications had been. Around half the participants also cited that they had based their sorting decisions on one type of statement involving recall of more information, a deeper level of processing, and reflecting more confidence than the other type of statement. The reasons that were endorsed by the fewest participants were those regarding use of more abstract words in one statement type and basing sorting decisions on gut instinct or guessing. On average participants selected 3.63 reasons (SD = 1.27) from the list of 10 (min. = 1, max. = 7).

Table 2.10. Decision criteria for statement sorting and the percentage of participants who endorsed each option.

Decision criterion	Percentage of participants who endorsed
More sure	67.3%
Recalled more information	55.1%
Deeper level of processing	53.1%
Confident	49.0%
Visual imagery	40.8%
Concrete words	36.7%
Emotional language	26.5%
Gut instinct	14.3%
Abstract words	12.2%
Mainly guessing	8.2%

Finally, participants were asked how confident they thought participants making the Type A and Type B statements had been and how much information or how many details they thought the participants had recalled in their memories for the cue words. Confidence was rated on a 0 to 100 percentage scale and information on a scale of 0 to 4. Using matching of Type A and B to Know and Familiar based on majority response, the mean confidence rating given for Know justifications was 60.27% (SD = 20.42) and for Familiar justifications was 56.53% (SD = 22.53), this difference was not significant, $t < 1$. Whilst the sorting criterion endorsed most by participants was differences in certainty (‘More sure’) between statement types, this was not reflected in the confidence values they assigned to the two types of justifications. However, ratings of how detailed the recall involved in the memory justifications had been did demonstrate a significant difference between Know and Familiar justifications. Familiar justifications were rated as containing more information and details than Know justifications, means 1.96 (SD = .84) and 1.47 (SD = .87) respectively, $t(48) = 2.24$, $p = .03$. 'Recalled more information' was the decision criterion with second highest endorsement by participants and this was reflected in their ratings of how detailed the Know and Familiar justifications were. This finding was explored further in three additional analyses.

Firstly, Know and Familiar justifications were compared on number of characters to see whether the higher level of information or details perceived in Familiar justifications might be due to the fact that they were longer. No difference in justification length was observed. Mean number of characters was 86.28 (SD = 33.57) for Know justifications and 91.56 (SD = 43.57) for Familiar justifications; $t(78.79) = 0.64$, $p = .52$. Secondly, it was examined whether ratings of level of information in Know and Familiar justifications differed for participants who endorsed or did not endorse 'Recalled more information' as a criterion they had used to split the justifications. Mean ratings are shown in Table 2.11.

Table 2.11. Mean information ratings (and standard deviations) split by whether the participant had endorsed 'Recalled more information' as a decision criterion.

	Endorsed 'Recalled more information'	
Justification type	Yes	No
Know	1.26 (.98)	1.73 (.63)
Familiar	2.15 (.82)	1.73 (.82)

A 2(endorsement) x 2(justification type) ANOVA was conducted with endorsement as a between-subjects factor and justification type as a within-subjects factor. ANOVA demonstrated a significant main effect of justification type, $F(1,47) = 4.39$, $p = .04$, and a significant interaction between justification type and endorsement, $F(1,47) = 4.39$, $p = .04$. No significant main effect of endorsement was observed, $F < 1$. Participants who had not endorsed 'Recalled more information' as a criterion they had used to sort the justifications did not rate Know and Familiar justifications as containing different levels of information or amount of details, $t < 1$. Conversely, participants who reported that they did base some of their sorting decisions on 'Recalled more Information' rated Familiar justifications as containing more information or details than Know justifications, $t(26) = 2.84$, $p = .009$. This interaction qualifies the significant main effect of justification type obtained from the ANOVA and the earlier t-test. While Familiar justifications were not seen as containing more information and details than Know justifications by all participants, for those participants who did identify differences in level of information as a decision criterion, these participants rated Familiar justifications as containing more information and details than Know justifications[10].

Finally, correlations were performed to examine whether there was a relationship between confidence ratings and ratings of level of information contained in the justifications. For both Know and Familiar justifications, correlational analysis demonstrated a significant positive relationship between confidence and information ratings, Know: $r(49) = .41$, $p = .003$, Familiar: $r(49) = .52$, $p < .001$. For both types of justification, the higher the amount of information and details the participant considered were reported in the memory justifications the more confidence they thought was reflected in the justifications. This demonstrates an element of construct validity between the ratings and suggests that, although participants' confidence ratings themselves were not significantly different for Know and Familiar justifications, there was a meaningful relationship between their interpretations of the confidence and information in the justifications.

2.5.4. Discussion

The central aim of this experiment was to examine whether participants were reliably able to identify two types of statement within the original Know justifications. Participants were not asked to categorise statements as Know or Familiar based on definitions, instead they were simply asked to split the statements into two types – Type A and Type B. How Type A and B mapped onto Know and Familiar was analysed in different ways. Using gamma correlation as a measure of association, results demonstrated that how participants split the justifications into Know and Familiar was reliably associated with expert classifications of Know and Familiar statements. Further analysis demonstrated that this association was stronger if mapping of Type A and B statement types was based on majority response rather than how participants reported they had mapped the statements on the

[10] A corresponding ANOVA comparing confidence ratings given to Know and Familiar justifications by participants who had endorsed 'More sure' as a decision criterion did not demonstrate any significant main effects of endorsement or justification type and no significant interaction (all $p > .25$).

questionnaire. This was due to some participants' questionnaire reports having no objective truth to them when compared to their sorting responses.

In addition to gamma correlations, the proportion of Know and Familiar statements that were correctly categorised was analysed. Know justifications were found to be categorised more reliably than Familiar justifications. Participants reported using amount of detail and certainty as the criteria they had most used to divide the statements into two types. Although there was no reliable difference between the confidence ratings assigned to the two statement types, participants did rate Familiar justifications as containing more information or details than Know justifications. In addition, further exploration of this difference revealed that it was only those participants who had endorsed amount of information or details as a decision criterion who rated Familiar justifications as containing more information than Know justifications. Participants had insight into their decision making processes and their endorsement and ratings on the decision criteria questionnaire reflected this.

The finding that Know justifications were considered to contain less information and fewer details than Familiar justifications provides a further element of support to the validity of separating these types of subjective experience. Gardiner et al. (1998) noted that their participants' Know justifications (which included all Know and Familiar justifications that were separated in the current experiment) were shorter and used rather limited vocabulary compared to Remember justifications. In the current experiment, Know and Familiar justifications were not found to differ in number of characters, so it was not length of statement which led participants to rate Familiar justifications as containing more information and details than Know justifications. Instead it is suggested that the type of memory processes that Gardiner et al.'s participants reported in their justifications is what led to these differences in ratings. Gardiner et al. indentified the original Know justifications as being absent of recollection of specific contextual details and instead being characterised by feelings of familiarity, just knowing, thinking a word occurred, or reporting of the absence of recollective details. It is suggested that it is aspects of these memory processes which differed when the original Know justifications were separated into Know and Familiar by expert raters CM and HLW. Participants in the current experiment appear to have identified the same differences, also used these to split the statements into two types, and given Know and Familiar justifications different ratings based on their own assessment of the underlying memory processes. Comprehensive linguistic analysis could be undertaken to further analyse the differences between the justification types and in future work it would be valuable to examine the justifications provided when participants are able to use all four response options of Remember, Know, Familiar, and Guess in the same manner which Gardiner et al. used to examine those for Remember, Know, and Guess.

In the current experiment the central analyses of interest demonstrated that participants were able to split Know justifications into separate Know and Familiar categories. Gamma correlations demonstrated that participants' categorisation of justifications as Know and Familiar was reliably associated with expert categorisation and the proportion of Know justifications categorised as Know was also found to be reliable. However, categorisation of Familiar justifications was not reliable; instead of consistently assigning Familiar justifications to Familiar, participants instead categorised some Familiar justifications as Know (as it was forced-choice; participants were simply splitting the justifications into two types). This is the opposite type of inappropriate categorisation to that observed in Experiments 2.2 and 2.3. In those experiments it was Know justifications which were often categorised as Familiar, particularly those Know justifications associated with low confidence (Experiment 2.2). The current experiment involved a very different task to those two experiments however. In Experiments 2.2 and 2.3 participants were provided with definitions of Remember, Know, Familiar and Guess and were asked to assign justifications to one type of subjective experience. In the current experiment participants were simply given a set of justification statements and were asked to sort them into two types; no further information was provided regarding types of subjective experience or how the statements should be sorted. It is suggested that it is these differences which led to the different patterns of categorisation of Know and Familiar justification statements across these three experiments. However, all three experiments found reliable differences between Know and Familiar justifications using a variety of different analyses. While patterns of results differed across the different tasks, together the results support the idea that Know and Familiar justifications reflect different underlying subjective experiences.

One surprising result in the current experiment was that there was no difference in confidence ratings for Know and Familiar justifications, as would be predicted from previous literature (e.g., Brewer & Sampaio, 2006; Dewhurst et al., 2009; Kihlstrom et al., 1996; Experiment 2.1 this chapter). If this finding is considered alongside the fact that categorisation of Familiar justifications was unreliable whereby they were quite often categorised inappropriately as Know, it could have been that the problems with categorisation influenced the confidence ratings and that performance for both was influenced by how difficult the task seemed. Participants reported that they found the two statement types very similar and the task was difficult. This is reflected in the fact that 14 participants' questionnaire reports of how they had mapped Type A and B to Know and Familiar did not match their actual sorting responses and eight participants did not provide an appropriate response to this question. Although there were limits to participants' awareness of the differences between categories of subjective experience in this difficult task, results indicate that, on the whole, people were able to distinguish two types of subjective experience from justifications that were all originally Know justifications. This further validates the separation of Know and Familiar as response options in experimental tasks.

2.6. General discussion

2.6.1. The split of Know into Know and Familiar

Results demonstrate that Know and Familiar can be reliably differentiated by lay people both in terms of inherent differences in the justification statements (Experiment 2.4) and in terms of the confidence associated with the justification (Experiment 2.1). The finding that Know and Familiar are associated with different levels of confidence extends previous findings observed in recognition memory paradigms to understandings of how lay people interpret reports of others' memory experiences (Brewer & Sampaio, 2006; Dewhurst et al., 2009; and Kihlstrom et al., 1996). Furthermore, Experiment 2.4 demonstrates that even when participants are not told how experts conceive of the differences between justifications, lay persons' interpretations of intrinsic differences are evident from their sorting responses and ratings of how detailed Know and Familiar justifications are.

However, the confidence associated with Know judgments was found to influence how they were categorised. In Experiment 2.2 a large proportion of Know justifications were assigned to Familiar if they were Medium or Low confidence justifications and in Experiment 2.3 medium-confidence Know justifications paired with any level of confidence were equally likely to be assigned to Familiar as to Know. As discussed in Section 2.4.4, it is suggested that this categorisation stems from the high confidence without recollection conceived as the basis of Know subjective experiences. Although these findings could suggest that the division of the Know category into separate categories of Know and Familiar is not valid, results from all experiments in this chapter have demonstrated different patterns of data for Know and Familiar justifications. Experiment 2.1 demonstrated significantly different confidence ratings for Know and Familiar justifications. Experiment 2.4 demonstrated that participants were able to reliably separate Know and Familiar justifications into two types and Familiar justifications were rated as containing significantly more information and details than Know justifications. Experiments 2.2 and 2.3 demonstrated that while Familiar justifications were consistently categorised as Familiar and not influenced by associated confidence level, Know justifications were. These different patterns of data support the idea that Know and Familiar subjective experiences are dissociable and validates the use of the four separate categories of subjective experience in episodic memory paradigms. All further experiments in this thesis asked participants to assign recognition responses to Remember, Know, Familiar, or Guess.

In sum, whereas the subjective experience of Familiarity appears to reflect the process of familiarity, the subjective state of Knowing appears to reflect high confidence in recognition without recollection.

2.6.2. Confidence and subjective experience

Experiment 2.1 demonstrated that mean confidence assigned to justification statements originally made for Remember, Know, Familiar and Guess responses differed significantly. Remember justifications were assigned higher confidence ratings than Know justifications, Know were higher than Familiar, Familiar higher than Guess. This demonstrates that lay people understand the relationship between subjective experience and confidence from others' memory experiences in the same way as is observed in experimental paradigms (Gardiner & Java, 1990; Dewhurst et al., 2009; Rajaram, 1993; Rajaram et al., 2002; Rotello & Zeng, 2008; Tulving, 1985; and Yonelinas, 2001a). Extending this, Experiment 2.2 demonstrated that when the confidence associated with a justification is manipulated intrinsically to the statement (High, Medium, and Low confidence statements selected as stimuli) this influenced how these justifications were assigned to subjective experience categories. Remember and Familiar justifications were demonstrated to be little-influenced by associated confidence level, whereas Know and Guess justifications were often categorised according to confidence. This suggests that the qualitative nature of Remember and Familiar justifications are not sensitive to manipulations of confidence while Know and Guess are.

Conversely, Experiment 2.3 demonstrated that when confidence was manipulated externally to the justification (a confidence value was provided alongside the statement) this had no influence over participants' categorisation of justifications. In Experiment 2.3 justifications were selected as 'prototypical' statements from around the mean level of confidence for each category of subjective experience and were paired with confidence values either appropriate to their original category or appropriate to another category of subjective experience. At all levels of confidence, Remember and Familiar justifications were again found to be consistently assigned to Remember and Familiar, while whatever confidence value was presented alongside Know and Guess justifications a large proportion of justifications were assigned to Familiar, paralleling the results obtained for Medium confidence justifications in Experiment 2.2. Participants did not take into account the confidence value presented alongside the justification statement. Across all levels of manipulated confidence, how Remember, Know, Familiar, and Guess justifications were assigned to categories of subjective experience remained consistent.

Given these findings, it seems difficult to suggest that confidence is merely the driving force behind judgments of subjective experience (Dunn, 2008). Instead it is suggested that people use experiential state to derive confidence. The importance of this was stressed by Tulving (1985) when he suggested that "the adaptive value of episodic memory and autonoetic consciousness lies in the heightened subjective certainty with which organisms endowed with such memory and consciousness believe, and are willing to act upon, information retrieved from memory" (p. 10). The current experiments have demonstrated that in understanding memory, reliable confidence values can be derived from subjective experiences, even when participants have not been involved in the original memory task.

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

The experiments presented in this chapter examined how different types of subjective experience are understood by people and found that confidence, confusion over source and familiarity, and level of detail involved in recall of a memory are all related to understanding of experiential state. These themes are carried forward in Chapters 3 and 4 of this thesis. In the experiments presented in Chapter 3, the familiarity of items is manipulated prior to study to explore how participants use different subjective experiences to overcome the confusion caused when all items are familiar. In Chapter 4, confidence, source, and subjective experience judgments are compared experimentally to examine how these are related for items recognised with and without recollection.

3. The Use of Recollection When Items are Familiar

"…estimates of recollection do not directly reflect the amount of qualitative information retrieved during item presentation, but rather the usefulness of this information in terms of selective responding."

(Dobbins, Kroll, Yonelinas, et al., 1998, p. 395)

In this chapter three experiments are presented where experimental familiarity of target and lure items was manipulated. Across three experimental designs and three types of recognition test, pre-exposure of items prior to study was found to impair memory performance. However, different patterns of subjective experience across pre-exposed and non pre-exposed items were only observed when a between-subjects design was used. The importance of experimental design when exploring subjective experience is discussed, as is the relationship between memory performance and subjective experience.

In Chapter 2, confusion over the source of familiarity for items was suggested to influence how justification statements were assigned to categories of subjective experience and levels of confidence. Confusion over the source of familiarity was manipulated experimentally in the current chapter, and the relationship between source, subjective experience and confidence were further examined in Chapter 4.

3.1. General introduction

Received wisdom suggests that the more familiar you are with material the better you will remember it. However, actually it is items that are more novel and distinctive (and therefore less familiar) that have been shown to be recognised better than items which are less distinctive (e.g., Gregg, 1976; Kinsbourne & George, 1974; Tulving & Kroll, 1995). How familiarity influences memory performance has been studied using two main paradigms: word frequency – using the frequency of a word's occurrence in the language as a variable; and experimental familiarity – where participants are pre-exposed to (some) items prior to study so those items increase in familiarity.

3.1.1. Word frequency effects

Word frequency effects are consistently found whereby words that are more frequent in the language, and therefore more familiar[11], e.g., sausage, are more easily *recalled*, but less frequent words, e.g., soufflé, are more easily *recognised,* with greater hit rates and lower false alarm rates than high-frequency words (see Gregg, 1976, and Gillund & Shiffrin, 1984, for reviews). However, for recall the pattern of results is not as consistent as it is for recognition. In recall, the design of the experiment has shown to be crucially important, with the advantage for high-frequency words being obtained when participants study pure lists consisting of either all high-frequency or all low-frequency words in a between-subjects design, but often being reversed or removed when the design is within-subjects using mixed lists (e.g., DeLosh & McDaniel, 1996; Gillund & Shiffrin, 1984; Gregg, Montgomery, & Castano, 1980).

In contrast, the advantage for less frequent words in recognition is more consistent and has been obtained using item recognition (e.g., Balota & Neely, 1980), forced-choice recognition (e.g., Glanzer & Bowles, 1976), and when the dependent measure is response speed rather memory performance (Duchek & Neely, 1989). Although in associative recognition, where the task is to recognise whether a test pair is an intact pair of words that were studied together or a recombination of words from other studied pairs, effects of word frequency are not so clear-cut. Clark (1992; Clark & Burchett, 1994) found an advantage for high-frequency words in associative recognition whereas Hockley (1984) found no effect of word frequency. These associative recognition results are therefore more consistent with the word frequency effects found for recall than for recognition. The issue of associative compared to item recognition will be returned to later (Section 3.2.1).

3.1.2. Word frequency effects and subjective experience

Overall, low-frequency words have an advantage when it comes to recognition. Several theories posit that the lower familiarity of these words is the basis for this advantage.

---

[11] However, frequency and familiarity differ subtly. Frequency refers directly to how often a word occurs in written and spoken language whereas familiarity refers to subjective ratings of how often a word is seen or experienced.

Mandler (1980) and Macleod and Kampe (1996) assert that the pre-experimental familiarity of a word determines how much of a 'boost' in familiarity it receives during study. As low-frequency words receive a larger 'boost' relative to high-frequency words, due to, for example, differential attending or increased rehearsal (Dobbins, Kroll, Yonelinas, et al., 1998), this leads to higher rates of recognition. Alternatively, models which focus on perceptual fluency as a basis for recognition, e.g., Jacoby and Dallas (1981), suggest that feelings of familiarity arise because of fluency and thus words that are relatively easy to perceive are experienced as familiar. For low-frequency words, because these words are more distinctive they will more often be recognised. Both Jacoby and Dallas (1981) and Mandler (1980) suggest that familiarity dominates in tests of recognition and that recollection is only relevant in certain situations such as recall where retrieval must be self-initiated (Mandler, 1980). However, studies which have investigated the effects of both recollection and familiarity with word frequency have shown that the frequency effect is mainly found in recollection-based recognition as opposed to familiarity-based recognition. To give one example, using the Remember-Know paradigm Gardiner and Java (1990) found that the low-frequency word advantage was only found in responses assigned to Remember, whereas equal proportions of high- and low-frequency words were assigned to Know.

Many other studies have since obtained similar findings. In his review of recollection and familiarity, Yonelinas (2002) concluded that the low-frequency advantage is almost always larger in recollection than in familiarity. Evidence comes from studies which have used the Remember-Know procedure (e.g., Cook, Marsh, & Hicks, 2005; Dewhurst et al., 1998; Gardiner & Java, 1990; Gardiner et al., 1997; Guttentag & Carroll, 1997; Huron, et al., 1995; Joordens & Hockley, 2000; Kinoshita, 1995; Reder, et al., 2000; Strack & Förster, 1995), the process-dissociation procedure (Guttentag & Carroll, 1997; Komatsu, Graf, & Uttl, 1995), and ROC studies (Arndt & Reder, 2002). Taken together, these studies demonstrate a low-frequency advantage for recollection and a smaller but reliable advantage for familiarity. Yonelinas (2002) calculated that low- compared to high-frequency words led to .16 and .09 mean increases in recollection and familiarity, respectively. Supporting evidence also comes from electroencephalography (EEG) studies. For example, Rugg, Cox, Doyle, and Wells (1995) demonstrated that event-related potentials (ERPs) to low-frequency words exhibited large and reliable Old-New effects (changes in the waveforms evoked by Old compared to New stimuli) and that these effects were non-significant in the ERPs evoked by high-frequency words.

Reder et al. (2000) proposed and tested a model they called the Source of Activation Confusion (SAC) theory of memory, which explains the word-frequency effect in terms of contributions of both recollection and familiarity. In the model, every word is considered to have a word node, information related to its pre-experimental familiarity; and an event node, information related to whether it has been encountered in the experiment. Activation can spread between nodes depending on the strength and amount of associations between nodes (Cary & Reder, 2003). Increases in false alarms for high-frequency words result from higher base activation of the word node of these items; participants are more inclined to spuriously accept high-frequency words as old because they have misattributed their pre-experimental familiarity. According to the model these words are therefore likely to be given a Know response (Cary & Reder, 2003; Reder et al., 2000). On the other hand, increases in correct recognition for low-frequency words occur because low-frequency words have fewer prior contextual associations competing with the current contextual association, more activation is able to reach their event nodes, making it more likely that the event node will pass over threshold and receive a Remember response. In summary, the studies discussed here have shown an advantage for less familiar, low-frequency words when familiarity is a facet of the word itself; however, manipulations of familiarity, or increases in competing contextual associations, have also been performed experimentally.

3.1.3. Experimental manipulations of familiarity

Manipulations of familiarity within an experiment have also demonstrated recognition advantages for less familiar items. An early study by Kinsbourne and George (1974) manipulated experimental familiarity by asking participants to complete a pre-exposure task on items where they rated the concreteness of the words. Half these pre-exposed items were then involved in the study-test procedure along with an equal number of novel words. Word-frequency was also examined with different lists containing high- and low-frequency words. This study found the usual low-frequency advantage in memory performance but also revealed an effect of experimental familiarity such that pre-exposure to words prior to study impaired recognition for both the low- and high-frequency words.

Kinsbourne and George's findings have been extended by a number of researchers over the last two decades (e.g., Åberg & Nilsson, 2001, 2003; Chalmers & Humphreys, 1998; Greene, 1999; Kormi-Nouri, Nilsson, & Ohta, 2005; Maddox & Estes, 1997; Tulving & Kroll, 1995). All studies concur that recognition accuracy is impaired for items that are pre-exposed prior to study. However, differences in source discriminability between pre-exposure and study tasks, and whether pre-exposure imposes its influence in correct recognition and/or false alarm rates has led to different theoretical interpretations from a number of studies.

Tulving and Kroll (1995) replicated Kinsbourne and George (1974) with the addition that participants were asked to rate their confidence in their responses. As well as finding that pre-exposed words were less well recognised, participants were also more confident in their correct recognition or correct rejection of words that had not undergone pre-exposure, compared to their confidence in their recognition and rejection of words that had been pre-exposed. Tulving and Kroll (1995) interpreted these findings in terms of their 'novelty-encoding hypothesis' which they derived largely from neuroimaging data regarding the existence of neurons which respond particularly to novelty (e.g., Tulving, Markowitsch, Craik, Habib, & Houle, 1996; Tulving, Markowitsch, Kapur, Habib, & Houle, 1994). This hypothesis posits that encoding proficiency is directly related to novelty within the experiment at the time of study, i.e., items that have been pre-exposed during the experiment will not be encoded as well during study as items that appear for the first time in the study phase.

Dobbins, Kroll, Yonelinas, et al. (1998) used a similar experimental design but included two different levels of pre-exposure. Tulving and Kroll (1995) had pre-exposed participants to each item six times and used the same semantic orienting task (living/non-living) during pre-exposure and study. These extensive pre-exposure manipulations led to the predicted finding that novel (not pre-exposed) studied items were correctly recognised to a greater extent than familiar (pre-exposed) studied items, but in addition nearly 50% of the pre-exposed but not studied items were falsely recognised. To explore whether the effects of pre-exposure on correct recognition would hold true when pre-exposure was not so extensive or whether the effect was mainly linked to the high amount of false recognition, Dobbins, Kroll, Yonelinas, et al. (1998) varied pre-exposure to one, two, or five preexposures. In addition, to create a situation where recollection could also play a role in overcoming the familiarity induced by pre-exposure they varied whether the same or different orienting tasks were used at pre-exposure and study (living/non-living or pleasant/non-pleasant).

Dobbins, Kroll, Yonelinas, et al. (1998) found a small overall effect of pre-exposure, with items being more likely to be recognised correctly if they had not been pre-exposed. However, performance interacted with whether the participant had undertaken the same or different orienting tasks at pre-exposure and study. Orienting task was manipulated between-subjects and for correct recognition the two groups did not differ but the same-task group produced more false alarms than the different-task group. Dobbins, Kroll, Yonelinas, et al. concluded that the main component of the difference in accuracy between the groups resulted from false alarms and that changing the orienting task improved recollection of items, making it easier for participants to discriminate between pre-exposed targets and distracters, overcoming the familiarity brought on by pre-exposure. These data speak against the novelty-encoding hypothesis being the explanation for the effects of pre-exposure. One of the key tenets of the novelty-encoding hypothesis is that the novelty assessment system screens out familiar items from further processing, however if the effect of pre-exposure is only to increase false alarms, then novelty of items at encoding has not helped improve recognition performance (Åberg & Nilsson, 2003; Dobbins, Kroll, Yonelinas, et al., 1998).

Whether source discriminability was the main factor influencing recognition performance to pre-exposed and non pre-exposed items was the focus of Åberg and Nilsson (2001) and Kormi-Nouri et al.'s (2005) extensions of this paradigm. These authors set out to increase the differences between the pre-exposure phase and the study phase in their experiments to try to test whether the different recognition performance was due to participants not being able to determine whether an item had been experienced only in pre-exposure or had also been studied, or it could be interpreted via the novelty-encoding hypothesis.

In two experiments by Kormi-Nouri et al. (2005) pre-exposure took the form of a, seemingly separate, subject-performed task (SPT) experiment (e.g., Cohen, 1981) where participants were presented with 90 commands (e.g., *'roll the pen'*). The verbs and nouns contained in these commands then became the materials for the critical phase, along with matched novel items. The critical phase of the experiment took the form of an incidental encoding task where participants were asked to make a frequency judgment about how often they performed the actions (e.g., *roll*) in daily life, or how often they used the objects (e.g., *pen*). After this incidental encoding task, participants underwent a recognition test for the words they had rated for frequency. The other words they had encountered in the SPT phase were used as familiar distractors and additional novel distractor words were also included. Participants were told to respond 'yes' to words they had rated for frequency, and 'no' to words they had not rated, however no specific instruction was given regarding recognition responses for items they had seen in the SPT phase.

Kormi-Nouri et al. (2005) found that, across their two experiments, pre-exposure to items led to impairments of recognition, or recognition superiority of novel items over familiar. This effect was found in recognition accuracy (hits minus false alarms), and in hits and false alarms analysed independently, contrary to Dobbins, Kroll, Yonelinas, et al. (1998) who found that the effect of pre-exposure was primarily due to increases in false alarms. Kormi-Nouri et al. found that pre-exposure reduced recognition for both types of material (verbs and nouns), in two languages (Swedish and Japanese), and for high- and low-frequency words, although it was more pronounced for verbs than for nouns, and for low-frequency words than for high-frequency words.

A further explanation is that difficulties with source discriminability are behind the effects of pre-exposure. Åberg and Nilsson (2001) used the same critical incidental encoding task as Kormi-Nouri et al., rating the frequency of occurrence of words, but used two different pre-exposure tasks; participants were either instructed to make up a short sentence containing the word or had to count the number of consonants in the word. However, the crucial difference in the study of Åberg and Nilsson was that in the final recognition test participants were explicitly instructed to only respond 'yes' to an item if they thought it had been on the study list (the incidental encoding list), regardless of whether they remembered it from the pre-exposure phase. Åberg and Nilsson found the pre-exposure effect for recognition accuracy data (hits minus false alarms) but not for hit rates alone, and suggested that the effect is due, at least in part, to source discrimination difficulties leading to increased false alarms in the pre-exposed items. However, Kormi-Nouri et al. (2005) suggested that the explicit test instruction used by Åberg and Nilsson may have led to them not finding an effect for hit rates because the explicit instruction increased problems of source discriminability rather than reducing it.

These studies have shown that when items are made familiar through pre-exposure prior to study recognition performance is impaired. Two hypotheses have been suggested for this finding. The ‘novelty-encoding hypothesis’ suggests that items which are encountered for the first time at study will be encoded more easily than items which are already familiar, resulting in better recognition performance for novel over familiar items. In contrast, the ‘source discriminability hypothesis’ proposes that it is problems of source confusion for familiar items which leads to more false alarms for these items and therefore superior recognition performance for novel items compared to familiar. The results of the current experiments are interpreted in terms of these two hypotheses in the General Discussion (Section 3.5.2).

3.1.4. Experimental manipulations of familiarity with measures of subjective experience

One aspect of the effects of pre-exposure/novelty on recognition that has been little researched is the subjective experiences on which recognition decisions are based when items have been pre-exposed. In their discussion of how explicit instructions to only report items from the critical study phase may have influenced recognition behaviour, Kormi-Nouri et al. (2005) state: “subjects had to think over Phases 1 and 2 and to decide whether they *remember* the information from the experiments or simply *know* the information” (Kormi-Nouri et al., 2005, p. 140, italics added). As reported above, Tulving and Kroll (1995) asked participants to make confidence judgments, with the finding that participants were more confident in their responses for words that had not undergone pre-exposure. In a later study, Åberg and Nilsson (2003) also had participants make confidence judgments to each item they recognised as being a studied item. Confidence judgments were made on a 3-point scale: 3 = very confident, 2 = more or less sure, and 1 = guessing; a scale which Åberg and Nilsson (2003) likened to that of Remember, Know, and Guess judgments of subjective experience. While the validity of this comparison is the subject of other chapters of this thesis (see in particular Chapter 4), Åberg and Nilsson (2003) did make some interpretation of their findings in relation to dual-process models and the Remember-Know paradigm.

In general, Åberg and Nilsson's (2003) results replicated the pre-exposure effects found by Tulving and Kroll (1995) but the effect was found to be down to increases in false alarms for pre-exposed items, not increases in hits for novel items. However, when analysis was performed on only those items given the highest confidence rating, an increase in correct recognition for novel items was found, which the authors took to be strong evidence against the source-discriminability hypothesis and for the novelty-encoding hypothesis. Åberg and Nilsson suggested that the reason a recognition advantage for novel items was only observed for responses with a high level of confidence may be because the novelty effect requires recollection of the encoding occasion. This finding parallels those observed in the word-frequency literature where the finding that low-frequency words are more easily recognised than high-frequency words has been shown to be obtained for Remember responses to a greater extent than for Know responses (e.g., Gardiner & Java, 1990; Yonelinas, 2002).

One paradigm which is similar to those involving pre-exposure is the process-dissociation procedure (PDP) of Jacoby (1991). As discussed in the General Introduction (Section 1.5.2.1) Jacoby (1991) developed the PDP to assess the contributions of recollection and familiarity to memory. In a PDP experiment participants study items from two lists (e.g., presented in different modalities) and then either recall items under inclusion conditions ('include all items regardless of list') – which is taken to measure the combination of familiarity and recollection processes; or under exclusion conditions ('only recall items from the second list') – taken to measure recollection. Recollection and familiarity are then calculated by comparing performance across these conditions. As mentioned above, findings of word-frequency effects using the PDP by Guttentag and Carroll (1997) and Komatsu et al. (1995) have found the same advantage for low-frequency words as those using the Remember-Know paradigm (Yonelinas, 2002).

Although the PDP may appear very similar to the experimental manipulations of familiarity via pre-exposure the critical difference is one of source discrimination. The main task in a PDP experiment is a source discrimination task – *where did I experience that item, was it List 1 or List 2?* This is similar to the source discrimination needed in the pre-exposure experiments discussed above, however Dobbins, Kroll, Yonelinas, et al. (1998) have pointed out that the inclusion and exclusion instructions are directly related to the qualitative processing undertaken at encoding (e.g., whether an item is spoken by a male or female voice) which is very different to the source discriminability present in many pre-exposure designs where typically either the pre-exposure phase or the study phase is one of incidental encoding and the instructions at recognition do not refer to the qualitative nature of the processing that took place at encoding. In line with this, many of the experiments discussed above were designed to try to increase the discriminability between pre-exposed and studied items, e.g., by changing the orienting task at pre-exposure and study (Åberg & Nilsson, 2001, 2003; Dobbins, Kroll, Yonelinas, et al., 1998; Kormi-Nouri et al., 2005). The current experiments also employ different tasks at pre-exposure and study with the prediction that this should produce a situation where recollection can play a role in overcoming the familiarity induced by pre-exposure.

The role of subjective experience in overcoming familiarity has been explored by Whittlesea and colleagues (e.g., Kronlund & Whittlesea, 2005, 2006; Whittlesea, 2002, 2004; Whittlesea & Williams, 2000, 2001a, 2001b). Their standard paradigm uses a sentence stem completion task, which is modified to induce a feeling of familiarity by the inclusion of a pause (on some trials) before the final word of the sentence appears. Using this paradigm, Whittlesea (2002) found that including a pause between sentence stem and terminal word led to greater hits and false alarms following high-constraint stems, for example, “She swept the floor with the… (pause) BROOM”. High-constraint stems are those for which only a few words would be able to complete the sentence sensibly (e.g., MOP, BRUSH). Whittlesea suggested that the pause in the sentence allowed participants to develop an expectation about the word that would complete the sentence. This expectation then invoked a feeling of familiarity with the word when it appeared, leading them to endorse the item as one they had seen before, even if they actually had not. A feeling of familiarity would not be induced in a low-constraint sentence such as “She couldn’t find a place to put the…”, as there is no expectation of what word will complete the sentence. Whittlesea (2002; Experiments 6a-c) went on to explore subjective experience responses in this paradigm by asking participants to report recognition decisions as either recall or familiarity.

In Experiment 6a, participants studied single words and were then presented with sentence stems at test; half the sentences were completed by the studied words and half were completely novel sentences. Here the feeling of familiarity induced by the pause between sentence stem and word at test led to more claims of recall in both hits and false alarms for high-constraint compared to low-constraint sentences. For high-constraint sentences, the feeling of familiarity induced by the pause led participants to claim they recalled the word from the study phase, even when it was a novel sentence. There were no differences in reports of familiarity.

In Experiment 6c, Whittlesea altered the procedure so that, at study, some words were studied as single words and others were studied in complete sentences. Later, at test, some words were presented with sentence stems that had been encountered with the word at study, and some were presented with novel sentence stems. With this procedure, both new and old high-constraint sentences were more likely to be claimed to be familiar than were low-constraint sentences. Here the feeling of familiarity induced when novel high-constraint sentence stems were presented led participants to endorse these new items as being familiar, not as being recalled as was found in Experiment 6a. Whittlesea explained these different findings by highlighting the differences in how items were experienced at test. In Experiment 6a all the sentence stems presented at test were novel, only single words had been presented at study. In contrast, in Experiment 6c, some of the stems had been encountered at study. In this experiment, encountering some pre-familiarised stems at test resulted in novel stems being claimed as familiar, not recalled. Thus, in these experiments the subjective experiences reported by participants depended not only on what they experienced for that individual item or sentence, they also depended on the nature of the other sentences included on the test (Whittlesea, 2002).

In contrast to the experiments which manipulated the *actual* familiarity of items experimentally (e.g., Åberg & Nilsson, 2001, 2003; Dobbins, Kroll, Yonelinas, et al., 1998; Kormi-Nouri et al., 2005), using his sentence completion paradigm with a pause between stem and terminal word Whittlesea (2002) was able to manipulate the *feelings* of familiarity experienced on particular trials and explore how these were reflected in judgments of subjective experience (see also Kronlund & Whittlesea, 2005, 2006; Whittlesea, 2004; Whittlesea & Williams, 2001b). Furthermore, Whittlesea demonstrated that when feelings of familiarity are manipulated in this way the subjective experience with which an item is recognised can depend on how that item is perceived in relation to the other items on the test. In view of that, it is suggested that in the current experiments, whether pre-exposure is manipulated within- or between-subjects could influence how recognition of items is experienced. Experiments that have examined subjective experience responses across experimental designs are discussed in the following section.

3.1.5. Experimental design for subjective experience experiments

One aspect of the experimental manipulations of familiarity not yet mentioned is that all the experiments discussed operationalised pre-exposure as a within-subjects variable. As the key area of interest is the subjective experiences that accompany decisions of recognition when items have been pre-exposed, it should be noted that, in addition to the Whittlesea (2002) findings discussed above, other manipulations have also demonstrated different patterns of subjective experience responses when carried out within- or between-subjects. For example, Dewhurst and Parry (2000) explored recognition for positive, negative, and neutral emotion words in both within- and between-subjects experimental designs. When mixed lists of words were presented (emotionality manipulated within-subject) positive and negative emotion words elicited more Remember responses than neutral words. However, this finding was eliminated when emotionality was manipulated between-subjects and participants studied pure lists of either positive, negative, or neutral words. Bodner and Lindsay (2003) also demonstrated the influence of test-list context on Remember-Know judgments in a levels-of-processing (LOP) experiment. They found that medium LOP items were more likely to be categorised as Remembered, and less likely to be classified as Known, when mixed with shallow rather than with deep LOP items, albeit overall recognition did not differ according to list context.

Another factor that can be problematic in within-subjects designs is carry-over from one condition to the other. Bodner and Richardson-Champion (2007) have recently explored carry-over effects and test-list context in an experiment on eyewitness testimony. They examined how subjective experience for recognition of details of a crime event were influenced by the relative difficulty of the other details on the test. Details of medium difficulty were found to be more likely to be classified as Remembered when mixed with hard details rather than easy details. In addition, when detail difficultly was operationalised in a blocked design, medium details were more likely to be classified as Remembered when preceded by a block of hard details rather than a block of easy details. However, informing participants of the relative difficulty of the upcoming block was found to eliminate the effect of blocking.

Similarly, Conway and Dewhurst (1995a) found, using a between-subjects design, that when trait words were encoded with reference to the self, words correctly recognised were more likely to be accompanied by a Remember response than when they were rated for valence at encoding. However, Hirshman and Lanning (1999) later attempted to replicate these findings using a within-subjects manipulation of encoding task but failed; instead they found equal levels of Remember responses across conditions. Conway and colleagues suspected that differences in the experimental design of their experiment and that of Hirshman and Lanning may have influenced the findings and Conway, Dewhurst, Pearson, and Sapute (2001) performed another series of experiments comparing experimental designs. These experiments demonstrated that the self-reference effect did replicate across between- and within-subjects designs but, in addition, found that results only replicated when the interval between study and test was 1 hour or longer, a difference in experimental design Hirshman and Lanning had not addressed.

Taken together, these studies demonstrate the importance of exploring subjective experiences using between-subjects experimental designs. As discussed by Conway et al. (2001), within-subjects designs can often be prone to one condition influencing another condition or ‘carry-over effects’. Experimental design for examination of subjective experience is critical to the three experiments presented in the current chapter. Here the effects of pre-exposure on subjective experience are examined using within-, between-, and blocked within-subjects designs.

3.1.6. Aims of the current experiments

The central aim of the three experiments presented here was to explore the subjective experiences underlying recognition decisions when items are pre-exposed prior to study. Experimental manipulations of familiarity have never before been investigated in conjunction with the Remember-Know paradigm. Here experimental familiarity and subjective experience were explored using between- (Experiment 3.1), within-subjects (Experiment 3.2), and a blocked within-subjects design (Experiment 3.3).

In line with previous findings regarding word frequency and experimental familiarity, the main prediction for each of these experiments was that pre-exposure to items prior to study would impair recognition performance. A second prediction was that pre-exposure would lead to higher levels of Remember judgments being made, as recollection can be used to determine which items were on the study list from those that were pre-exposed. It was also predicted that use of familiarity as a basis for recognition would lead to impaired memory performance for items that had been pre-exposed prior to study.

With regard to the source-discriminability and novelty-encoding hypotheses, if an effect of pre-exposure is demonstrated in Experiment 3.1, where pre-exposure is manipulated between-subjects, this result would provide support for the source-discriminability hypothesis. However, the novelty-encoding hypothesis cannot be formally tested in Experiment 3.1 as there are no differences in the novelty/familiarity of items within-subjects at encoding. In contrast, with the within-subjects manipulation of pre-exposure in Experiment 3.2 both the source-discriminability and the novelty-encoding hypotheses can be tested. Patterns of results for memory performance and subjective experience are discussed in relation to these two hypotheses in the General Discussion (Section 3.5.2).

3.2. EXPERIMENT 3.1: RECOLLECTION IN ASSOCIATIVE RECOGNITION

3.2.1. Introduction

In this experiment, participants studied forename-surname pairs. At recognition participants were then presented with the studied surname and two forenames and their task was to select the forename which had been studied with that surname. The use of associative encoding here is due to previous findings regarding use of recollection processes in different types of recognition. Yonelinas (1997, 2002) argues that associative recognition is primarily based on recollection, whereas item recognition is based to a greater extent on familiarity. In line with this, Hockley and Consoli (1999) have shown that associative recognition decisions are identified as Remember responses to a greater extent than are item recognition decisions and other authors have suggested that associative recognition decisions are based, at least in part, on a recall-like retrieval process (e.g., Humphreys, 1978; Clark, 1992; Clark & Burchett, 1994; see Clark & Gronlund, 1996, for a review). By this line of reasoning, one might expect pre-exposure not to interfere with associative recognition to the same extent as it would item recognition.

As discussed earlier, having different orienting tasks at pre-exposure and study should also increase the role of recollection in overcoming familiarity (Dobbins, Kroll, Yonelinas, et al., 1998). In the current experiment, the use of associative encoding at study and single items at pre-exposure was introduced to bring about the same effect.

3.2.2. Method

3.2.2.1. Participants and Design

Participants were 67 undergraduate students (59 female), mean age 19.01 years (range 18 to 22), from the University of Leeds, who received either Participant Pool Credits or £5 payment. Participants were tested in groups of between three and seven at individual PCs. The experiment employed a between-subjects design; in the Pre-Exposure condition (N = 34) participants undertook the pre-exposure familiarisation task prior to the study phase whereas in the No Pre-Exposure condition (N = 33) participants did not. Participants were assigned to conditions in a pseudo-random manner based on the timeslot they had signed-up for; conditions were equally distributed across the day.

3.2.2.2. Materials

The to-be-learnt items consisted of 40 high-frequency forename-surname pairs (e.g., Sophia Watson). Twenty female and twenty male forenames were taken from the Office of National Statistics (www.statistics.gov) as the most popular given names in England and Wales in 2007. The forenames were paired with the 40 most common surnames in England and Wales in 2007 taken from the National Health Service Central Register. Pairing was done pseudo-randomly, with matching initial letters (e.g., George Griffiths) avoided. Different spellings of the same name and unisex forenames (e.g., Alex) were excluded, and forename-like surnames (e.g., Thomas) were excluded as surnames.

A further 20 female and 20 male forenames were taken from the same website to be the lure forenames in the 2-alternate-forced-choice (2AFC) recognition test. Matching ensured the lure did not have the same initial letter as the target forename or the surname. Target and lure forenames were also matched so that there were approximately equal numbers (13 or 14) of target-lure pairs which did and did not match on gender (e.g., Thomas & Ryan, versus Eva & Matthew; all target-lure-surname triads are shown in Appendix B.2).

3.2.2.3. Procedure

In the Pre-Exposure condition, prior to studying the 40 forename-surname pairs, participants performed a separate pen and paper task – counting the number of vowels in each of the 120 names (40 surnames, 40 target forenames, and 40 lure forenames) with the aim of equating familiarity of the target and lure names. Participants in the No Pre-Exposure condition did not perform this task. The subsequent study phase and test phase procedures and instructions were identical for all participants. All instructions, study materials, and recognition test materials were presented, and data collected, using E-Prime version 1.2.

In the study phase, each of the 40 forename-surname pairs were presented, in a random order, individually on the computer screen for five seconds, separated by a fixation point presented for one second. Participants were instructed that they should study the name pairs and that later they would undergo a 2AFC recognition test where they would have to recognise which of two forenames had been studied with that surname. Prior to undergoing the test phase, participants completed 100 sums as a distracter task which took approximately 15 minutes.

In the test phase, participants were presented with instructions and given time to familiarise themselves with the definitions of Remember (R), Know (K), Familiar (F), and Guess (G) which they were to use to judge their recognition experience for all items (see Appendix A.2). Participants were then presented with the 40 surnames each accompanied by two forenames – the target name and a lure name, placed above one another to the left of the surname (presentation order randomised and placement on screen counterbalanced across items). Participants selected which of the two forenames they recognised as having been previously paired with that surname using the mouse and reaction time was recorded. For each item, the participant then made a subjective experience rating by clicking the R, K, F, or G box on the computer screen. When each recognition screen and subjective experience screen appeared the mouse pointer returned to the centre of the screen.

3.2.3. Results

Firstly, to ascertain whether pre-exposure to the names influenced memory performance, overall recognition performance by participants in the two conditions was examined. Proportions of correct responses assigned to each of the subjective experience categories (Remember, Know, Familiar, and Guess) were then calculated using both *a priori* and *a posteriori* methods; the memory performance of participants in each condition was correlated with assignment of items to the different categories of subjective experience; and finally, recognition decision reaction times were analysed by their later subjective experience category. Prior to analysis, data from any responses which had been made faster than 300ms or slower than 8000ms were excluded from the dataset.

3.2.3.1. Memory measures

The proportions of forenames correctly selected in the 2AFC recognition test were examined to assess whether there was a difference in memory performance between participants in the Pre-Exposure condition and the No Pre-Exposure condition. In both conditions mean recognition performance was found to be high, being .83 (SD = .11) in the Pre-Exposure condition and .87 (SD = .07) in the No Pre-Exposure condition. Because the strong prediction was made that pre-exposure would reduce performance, the strategy adopted was to use one-tailed tests to compare performance in the two conditions. Recognition performance was indeed found to be significantly higher in the No Pre-Exposure condition, $t(65) = 1.85$, $p = .034$ (one-tailed), indicating that prior exposure to experimental materials impaired memory performance. For the sake of completeness, recognition was also examined using *d'*. The *d'* values[12] in the Pre-Exposure and No Pre-Exposure conditions were 1.41 (SD = .60) and 1.66 (SD = .53) respectively, mirroring the significant difference in memory performance found using the raw recognition proportions, $t(65) = 1.80$, $p = .038$ (one-tailed).

3.2.3.2. Subjective experience measures at retrieval

The second analysis of interest was what subjective experiences had participants used to make their recognition decisions. Firstly *a priori* proportions were calculated – the proportion of correct responses assigned to each of the subjective experience categories.

[12] Formulas for *d'* are shown in Appendix C.

These proportions sum to 1 for each participant which can be problematic for statistical analysis as it violates assumptions of independence; however, following approaches described elsewhere (e.g., Conway et al., 1997; Herbert & Burt, 2004), where this occurs it is noted and the variables are analysed together with planned comparisons used to further explore the data. As is shown in Figure 3.1, correct recognition was based on different types of subjective experience in the two conditions.

Participants in the Pre-Exposure condition assigned .39 of correctly recognised names to the Remember category, and only .20 to the Know category, whilst participants in the No Pre-Exposure condition assigned .29 to Remember and .29 to Know. In both conditions, assignment of correctly recognised names to the Familiar and Guess categories were similar; .24 and .25 were assigned to Familiar in Pre-Exposure and No Pre-Exposure respectively and .17 were assigned to Guess in both conditions.

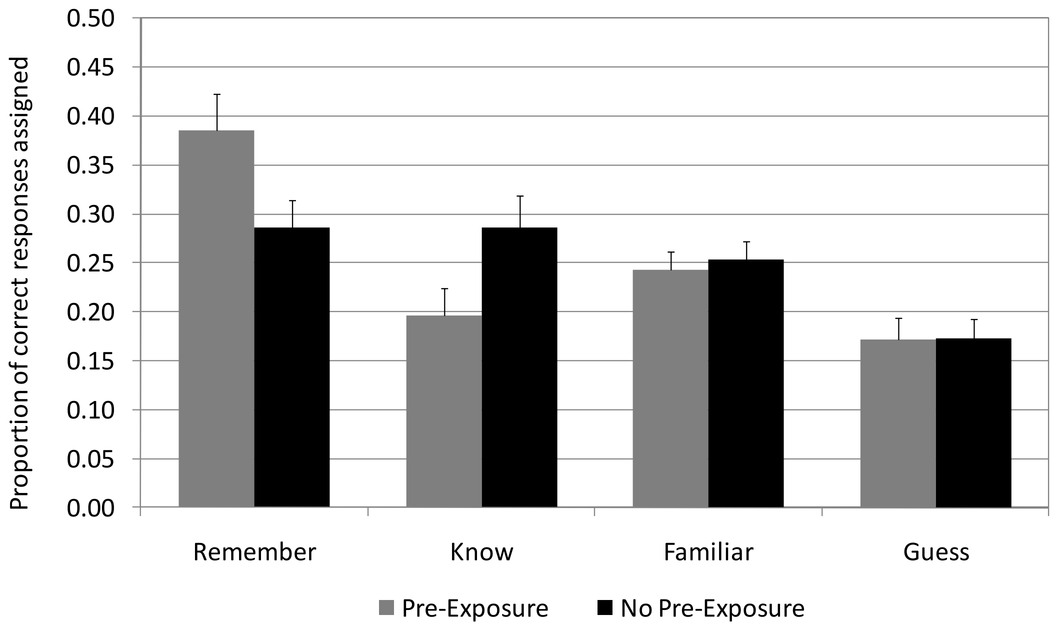

Figure 3.1. Proportion of correct responses assigned to Remember, Know, Familiar, and Guess by participants in the Pre-Exposure and No Pre-Exposure conditions. Errors bars = 1 SeM.

A 2(exposure condition) x 4(subjective experience) ANOVA demonstrated a significant main effect of subjective experience, $F(3,195) = 9.61$, $p < .001$, and a significant interaction between exposure condition and subjective experience, $F(3,195) = 3.18$, $p = .025$. No effect of exposure condition as a between-subjects factor was calculated as subjective experience categories summed to 1. Planned comparisons between the conditions found significantly more responses were assigned to Remember in the Pre-Exposure condition than in the No Pre-Exposure condition, $t(65) = 2.13$, $p = .037$. Conversely, significantly fewer responses were assigned to Know in the Pre-Exposure condition than in the No Pre-Exposure condition, $t(65) = 2.06$, $p = .044$. Familiar and Guess comparisons were non-significant, $t < 1$. Comparisons across subjective experience categories in the two conditions also revealed a significantly greater proportion of responses had been assigned to Remember as opposed to Know in the Pre-Exposure condition, $t(33) = 3.28$, $p = .002$, but not in the No Pre-Exposure condition, $t < 1$. In the Pre-Exposure condition proportion of responses assigned to Remember was significantly larger than proportion assigned to any other subjective experience category, all at least $p < .006$, and also significantly more responses had been assigned to Familiar than to Guess, $t(33) = 2.70$, $p = .01$. In contrast, in No Pre-Exposure no differences were found between the proportions assigned to Remember, Know, or Familiar, all $t < 1$; here only Guess was assigned a significantly lower proportion of responses than all other categories, all at least $p < .02$.

Since the critical hypothesis was that the Pre-Exposure condition would influence experiential states and in turn memory performance, *a posteriori* proportions were also of interest. These take into account the amount of incorrect responses as they calculate whether an item that had been assigned to a particular subjective experience category was more likely to have come from a correct or incorrect recognition response. For both exposure conditions, items were more likely to be correctly recognised if assigned to Remember and Know categories than to Familiar or Guess categories, see Figure 3.2. The proportion of responses that had been correctly recognised was over .90 for all Remember and Know responses. These were both higher than for Familiar responses, and finally Guess responses had the lowest likelihood of having been assigned to a correct response as opposed to assigned to an incorrect response.

A 2(exposure condition) x 4(subjective experience) ANOVA demonstrated a significant main effect of exposure condition as a between-subjects factor, $F(1,61) = 5.52$, $p = .022$, and a significant main effect of subjective experience, $F(3,183) = 57.07$, $p < .001$, but no interaction between exposure condition and subjective experience, $F < 1$. Planned comparisons within conditions revealed that in both the Pre-Exposure and No Pre-Exposure conditions there was no significant difference between the likelihood of recognition decisions having been correct when items were assigned to the Remember category compared to the Know category, both $t < 1$, whereas all other comparisons between categories were significant, all $p < .002$. The likelihood of recognition decisions being correct was equally high for responses which went on to be classified as Remember or Know, and for both conditions was significantly higher than the likelihood of items being correct which went on to be assigned to Familiar or Guess.

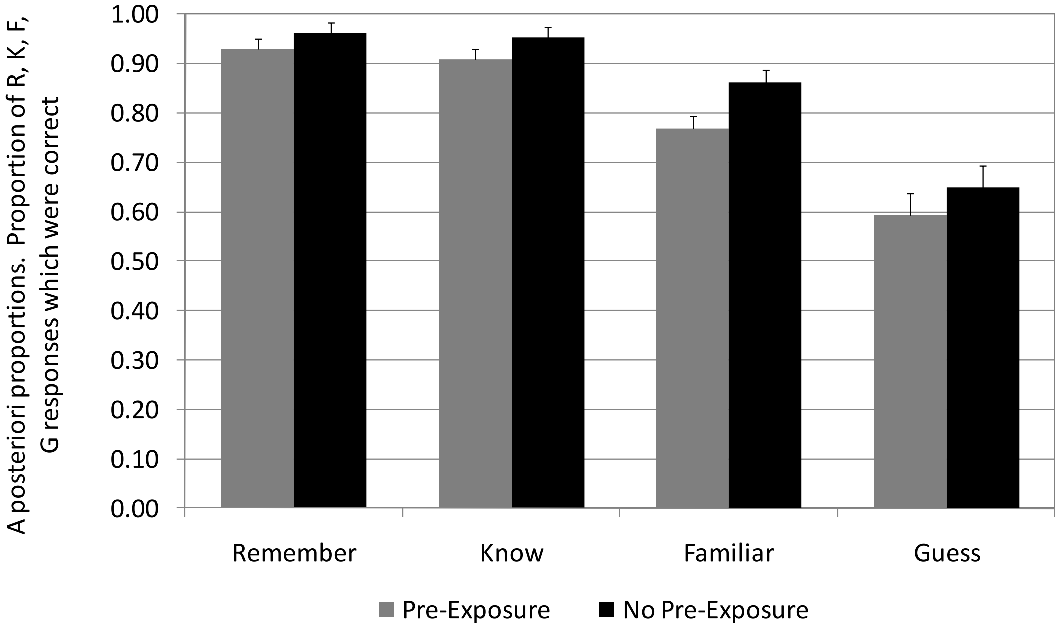

Figure 3.2. Proportion of forenames assigned to Remember, Know, Familiar, and Guess in the two conditions that had previously been correctly recognised. Errors bars = 1 SeM. N = 31 in Pre-Exposure condition and N = 32 in No Pre-Exposure condition due to missing data.

Comparisons between the conditions revealed that only for recognition decisions which were assigned to Familiar were responses in the No Pre-Exposure condition significantly more likely to have come from a correct response than responses in the Pre-Exposure condition, $t(65) = 2.46$, $p = .017$; all others $p > .15$.

3.2.3.3. Correlational analyses

To examine whether recognition based on a particular type of subjective experience was associated with improved or impaired performance, the proportion of names correctly recognised was correlated with the assignment of responses to the different levels of subjective experience; see Table 3.1. In the Pre-Exposure condition, memory performance was found to be significantly positively correlated with proportion of correct items assigned to Remember, and significantly negatively correlated to proportion of correct items assigned to Familiar. That is, people who made more Remember responses tended to have higher levels of performance and people who made more Familiar responses tended towards lower levels of performance. Correlations between performance and number of items assigned to Know, and Guess, were not significant. In the No Pre-Exposure condition, memory performance was not significantly correlated with proportion of names assigned to any of the levels of subjective experience.

Table 3.1. Correlations between memory accuracy and proportion of items assigned to subjective experience categories in each condition.

Condition	Remember	Know	Familiar	Guess
Pre-Exposure	.48**	-.07	-.57**	-.25
No Pre-Exposure	.17	-.01	-.01	-.21

N = 34 Pre-Exposure, 33 No Pre-Exposure. ** $p < .005$.

To test whether there were significant differences between the correlations of memory performance with use of subjective experience in the two conditions Fisher's Z was calculated using Fisher's r-to-z transformations of the correlation coefficients (Fisher, 1921; see Appendix D for full formulae). Although this does not allow for full exploration of the interactions between all coefficients, it does allow us to see whether key comparisons are significantly different. Comparing the coefficients for Remember, Know, and Guess across the two conditions revealed no significant differences, all $Z < 1.4$. However the negative correlation between memory performance and proportion of items assigned to Familiar in the Pre-Exposure condition was found to be significantly greater than that in the No Pre-Exposure condition, $Z = 2.49$, $p < .05$.

In summary, pre-exposure reduces correct responses, but increases the number of correct items assigned to Remember (as a proportion). An interpretation of this finding is that the experience of Remembering reflects recollection processes being used to disambiguate familiar stimuli. This is borne out in the correlational data – participants in the Pre-Exposure condition who reported more Remember responses had higher performance. Conversely, one of the strongest findings was that participants who made more Familiar responses had poorer performance. Furthermore, the negative correlation between familiarity and performance for participants who had been Pre-Exposed was significantly larger than that for participants who had No Pre-Exposure where there was no relationship between objective performance and reports of subjective experience.

3.2.3.4. Reaction times for recognition decisions

As discussed earlier (General Introduction, Section 1.5.1.1), a point of debate between dual-process and single-process theorists is whether there are differences in the time it takes to make recognition decisions based on different subjective experiences. Many dual-process models predict that responses based on familiarity should be quicker than those based on recollection as familiarity is considered to be a rapid and automatic process with relatively low demands on cognitive resources while recollection is a slower and more effortful process (e.g., Jacoby, 1991; Mandler, 1980; Yonelinas & Jacoby, 1994, 1996). However, many studies have shown recognition decisions for Remember responses to be reliably faster than for Know responses (e.g., Dewhurst & Conway, 1994; Dewhurst et al., 1998, 2006; Henson et al., 1999; Vilberg & Rugg, 2007; Wixted & Stretch, 2004). In the current experiment reaction times (RTs) for 2AFC correct recognition decisions were recorded and categorised by their later subjective experience judgment. As shown in Figure 3.3, mean RTs were faster for responses later categorised as Remember or Know than those categorised as Familiar or Guess.

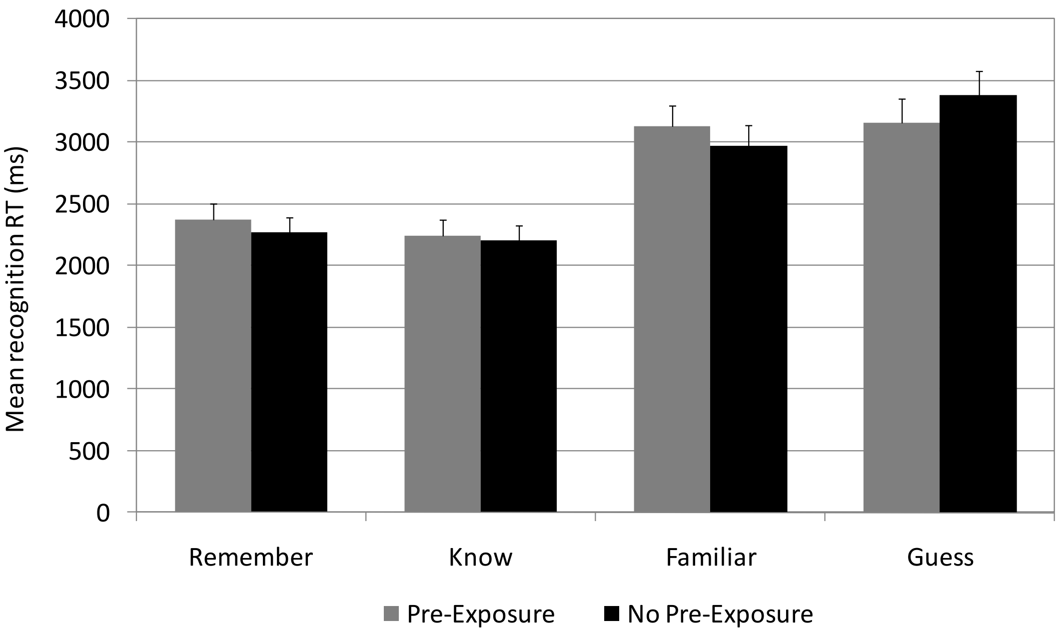

Figure 3.3. Mean correct recognition RTs by later subjective experience category. Error bars show 1 SeM. N = 30 in both conditions due to missing data.

A 2(exposure condition) x 4(subjective experience) ANOVA demonstrated no effect of exposure condition as a between-subjects factor, $F < 1$, a significant main effect of subjective experience, $F(3,174) = 47.74$, $p < .001$, and no interaction between exposure condition and subjective experience category, $F(3,174) = 1.31$, $p = .27$. Collapsed across group, comparisons demonstrated that the speed with which recognition decisions were made which were later classed as Remember or Know was not significantly different, $t(59) = 1.02$, $p = .31$. However Remember and Know decisions were significantly faster than both Familiar and Guess decisions, all $p < .001$; and finally, correct recognition decisions based on Familiarity were made significantly faster than Guesses, $t(59) = 2.33$, $p = .02$.

Thus, although memory is impaired by pre-exposure, it does not lead to a significant difference in reaction times. Both groups made Remember, Know, Familiar, and Guess responses equally according to the time taken to respond. Pre-exposure impairs performance and shifts responses to Remember, but it does not operate to shift reaction times. Reaction times show the same pattern in both conditions suggesting that subjective reports are reflective of the same underlying processes in each group.

3.2.4. Discussion

In this experiment the manipulation of pre-exposing participants in one condition to all the first names and surnames that would be encountered during the study and test phases of the experiment led to differences in all analyses. Overall, memory performance was found to be lower in the Pre-Exposure condition compared to the No Pre-Exposure condition which replicates the findings of Dobbins, Kroll, Yonelinas, et al. (1998), Kormi-Nouri et al. (2005), Åberg & Nilsson (2001, 2003), Tulving and Kroll (1995), Maddox and Estes (1997), Chalmers and Humphreys (1998), and Greene (1999).

The novel finding of the current experiment was that patterns of use of the subjective experience categories differed depending on whether items had been pre-exposed or not. More correctly-recognised names were assigned to Remember and less to Know in the Pre-Exposure condition compared to the No Pre-Exposure condition. Taken in conjunction with the memory performance results, this analysis shows that with Pre-Exposure, while use of recollection increases – as indicated by a greater proportion of Remember responses, memory performance actually decreases. This goes against the implicit assumption that the more recollection you have the better your memory performance.

Analysis of *a posteriori* proportions also demonstrated that when participants made their recognition decisions based on familiarity, this decision was more likely to be incorrect in the Pre-Exposure condition than the No Pre-Exposure condition; the familiarity induced by pre-exposure was an inappropriate cue for making a decision. This finding was reiterated by the negative correlation between familiarity and poorer memory performance in the Pre-Exposure condition. No relationships between subjective experience and memory performance were demonstrated in the No Pre-Exposure condition indicating that when items did not undergo pre-exposure, it was not necessary to use a particular type of subjective experience to aid recognition. Discussion of these findings with respect to the previous literature on confidence and recollection in experimental manipulations of familiarity (e.g., Tulving & Kroll, 1995; Åberg & Nilsson, 2003) and theoretical explanations associated with the novelty-encoding and source-discriminability hypotheses (e.g., Kormi-Nouri et al., 2005; Dobbins, Kroll, Yonelinas, et al., 1998) are left until the General Discussion (Section 3.5.2).

Reaction times in this experiment replicated previous findings, with recognition decisions based on recollection being made more quickly than those based on familiarity. However, in this experiment the use of separate Know and Familiar categories meant that the findings do not map neatly onto earlier Remember-Know data as here it was Remember and Familiar responses which demonstrated the difference, not Remember and Know responses. In the current experiment Know responses were made just as quickly as Remember responses. This issue is returned to in the General Discussion (Section 3.5.3).

Overall, these findings replicate those observed in previous experiments on word frequency and experimental manipulations of familiarity: memory performance was impaired for items pre-exposed prior to study. The novel finding of this experiment was that the differences in performance observed in conditions of pre-exposure and no pre-exposure can be understood by the underlying processes and subjective experiences on which recognition decisions are based in the different conditions. This experiment also demonstrated, for the first time, that pre-exposure can impair recognition when manipulated using a between-subjects design.

3.3. Experiment 3.2: Recollection when familiarity is manipulated within-subjects

3.3.1. Introduction

As all previous experiments manipulating pre-exposure had used within-subjects designs the aim of Experiment 3.2 was to explore whether the results from Experiment 3.1 would be replicated within-subjects with participants undergoing pre-exposure on half the stimuli and not on the other half. It was expected that a within-subjects manipulation would influence participants' use of the different categories of subjective experience, and that this could lead to changes in overall recognition performance. The methodology employed in this replication was also somewhat different to that of Experiment 3.1. Participants were recruited in lecture classes and therefore the study phase was presented on Powerpoint and all other tasks completed in a test booklet.

3.3.2. Method

3.3.2.1. Participants and Design

Participants were 44 students, 27 from the University of Leeds (22 female) and 17 from Åbo Akademi, Turku, Finland (16 female). The British students had a mean age of 23.26 years; 14 were third-year students and 13 were MSc students. The Finnish students came from a range of degree courses; their mean age was 21.24 years. The Finnish participants had been learning English for a minimum of six years (mean 8.24) and all rated their understanding of English as at least four on a scale of 1-7 (mean 5.59). All analyses were initially performed with nationality as a between-subjects variable and no differences were observed so this factor was removed. Participants were tested in their respective course groups in lecture theatres. This experiment employed a within-subjects design with participants being pre-exposed to only half the study stimuli.

3.3.2.2. Materials

The stimuli were identical to those in Experiment 3.1, however here the Pre-Exposure list of names included only half the target forenames, half the surnames and half the lure forenames (60 names total). The lure forenames which were pre-exposed were those which were later paired with those forenames and surnames also pre-exposed, i.e., for each 2AFC combination in the test phase participants had either been exposed to all elements of the 2AFC combination (the target forename, lure forename, and surname) or had not been pre-exposed to any of those names. The Pre-Exposure set of names was counterbalanced across participants.

A test booklet was created which included the consent form, demographics questionnaire, Pre-Exposure task, a short distracter task which consisted of a pattern comparison task (Salthouse, 1991), the definitions of R, K, F, and G, and finally the recognition test. Due to trying to limit the duration of the experimental session the distracter task was kept short and took participants approximately three minutes to complete. A Powerpoint slideshow was created to present the forename-surname pairs for study.

3.3.2.3. Procedure

Participants completed the demographic questionnaire and then performed the Pre-Exposure task – counting the number of vowels in the 60 names (20 surnames, 20 target forenames, and 20 lure forenames). When all participants had completed this task the Powerpoint slideshow was presented on the screen at the front of the lecture theatre.

In the study phase the 40 forename-surname pairs were presented individually; all participants studied the items in the same order. Participants were instructed that they should study the name pairs and that later they would undergo a 2AFC recognition test where they would have to recognise which of two forenames had been studied with that surname. Participants studied each name-pair for four seconds with a fixation point presented for one second between stimuli. Participants then returned to their test booklets and completed the distracter task, proceeding on to the recognition test in their booklets in their own time.

In the test booklets participants were presented with instructions and told to familiarise themselves with the definitions of Remember (R), Know (K), Familiar (F), and Guess (G) which they were to use to judge their recognition experience for all items; these definitions were identical to those used in Experiment 3.1 (see Appendix A.2). Participants were then presented with the 40 surnames each accompanied by two forenames – the target name and a lure name, placed above one another to the left of the surname. Participants were instructed to circle which of the two forenames they recognised as having been previously paired with that surname. For each item the participant was then instructed to circle either R, K, F, or G to make their subjective experience judgment.

3.3.3. Results

The analysis for Experiment 3.2 took the same approach as in Experiment 3.1. Firstly, to determine whether pre-exposure to the names influenced memory performance recognition accuracy for pre-exposed and non-pre-exposed names was examined. Proportions of responses assigned to each of the subjective experience categories (Remember, Know, Familiar, and Guess) were then calculated using *a priori* and *a posteriori* methods. Lastly memory performance of participants to pre-exposed and non-pre-exposed names was correlated with assignment of those items to the different categories of subjective experience.

3.3.3.1. Memory measures

To assess whether, using this within-subjects design, participant's memory performance was better for names to which they had not been pre-exposed the proportions of forenames correctly selected in the 2AFC recognition test in the two conditions were examined. In both conditions mean recognition performance was found to be high, being .81 (SD = .12) in the Pre-Exposure condition and .86 (SD = .10) in the No Pre-Exposure condition. As found in Experiment 3.1, recognition performance was significantly higher in the No Pre-Exposure condition than in the Pre-Exposure condition – participants were able to recognise more forenames correctly when they had not encountered them on the vowel-counting task prior to study, $t(43) = 3.65$, $p = .001$ (one-tailed). Recognition was also examined using *d'.* The *d'* values in the Pre-Exposure and No Pre-Exposure conditions were 1.12 (SD = .53) and 1.41 (SD = .56) respectively, mirroring the significant difference in memory performance found using the raw recognition proportions, $t(43) = 4.21$, $p < .001$ (one-tailed).

3.3.3.2. Subjective experience measures at retrieval

Both *a priori* and *a posteriori* proportions were calculated. In this experiment, proportion of correct recognition responses assigned to each of the categories of subjective experience did not differ between the two conditions; see Figure 3.4.

In comparison to the differing use of Remember and Know categories of subjective experience in Experiment 3.1, here similar proportions of correctly recognised names were assigned to the Remember category whether the name had undergone No Pre-Exposure (.37) or had been Pre-Exposed (.35). Additionally, similar levels were assigned to Know whether the name had undergone No Pre-Exposure or had been Pre-Exposed (.19 and .17 respectively). Here proportions for Remember and Know in both conditions are similar to the proportion of responses assigned to Remember and Know in the Pre-Exposure condition in Experiment 3.1 (.39 and .20 respectively) – use of subjective experience categories in both conditions mirrors the patterns observed in the Pre-Exposure condition in Experiment 3.1. Proportions of correctly recognised names assigned to Familiar and Guess categories were approximately equal in the two conditions as they had been in Experiment 3.1, however use of Familiarity as a basis for recognition decisions was increased and use of Knowing reduced in this experiment compared to in Experiment 3.1.

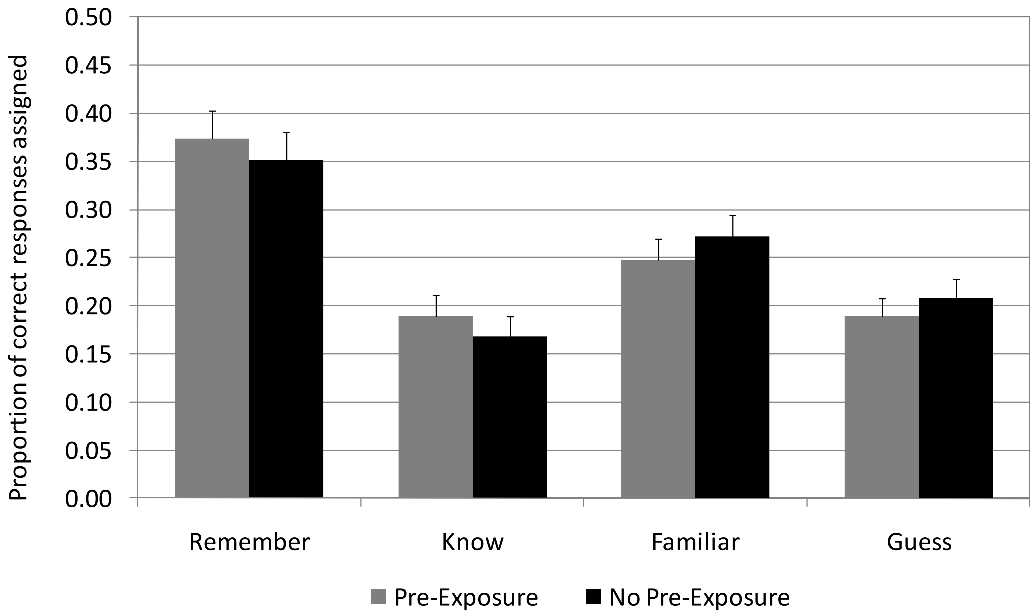

Figure 3.4. Proportion of correct responses assigned to Remember, Know, Familiar, and Guess by participants in the Pre-Exposure and No Pre-Exposure conditions. Errors bars = 1 SeM.

A 2(exposure condition) x 4(subjective experience) within-subjects ANOVA was conducted on these proportions. A significant main effect of subjective experience, $F(3,129) = 12.14$, $p < .001$ was observed, but no interaction between exposure condition and subjective experience level, $F < 1$. As subjective experience proportions summed to 1, an effect of exposure was not analysed, however comparisons between exposure conditions within each subjective experience category revealed no significant differences, all $p > .20$. Collapsed across conditions, t-tests demonstrated significantly more responses had been assigned to Remember than to any other category of subjective experience, all at least $p < .01$. A significantly greater proportion of responses had also been assigned to Familiar compared to Know, $t(43) = 2.95$, $p = .005$, and Familiar compared to Guess, $t(43) = 2.43$, $p = .02$. There was no significant difference between how many responses were assigned to Know and Guess, $t < 1$.

A posteriori probabilities were then calculated on these data – whether an item that had been assigned to a particular subjective experience category was more likely to have come from a correct or incorrect recognition response. From Figure 3.5 it can be seen that the pattern of means matches those observed in Experiment 3.1 – for each category of subjective experience, the probability of an item coming from a correct recognition response was lower for items in the Pre-Exposure condition than in the No Pre-Exposure condition. In addition, in both exposure conditions items were more likely to be correctly recognised if assigned to Remember and Know than to Familiar or Guess.

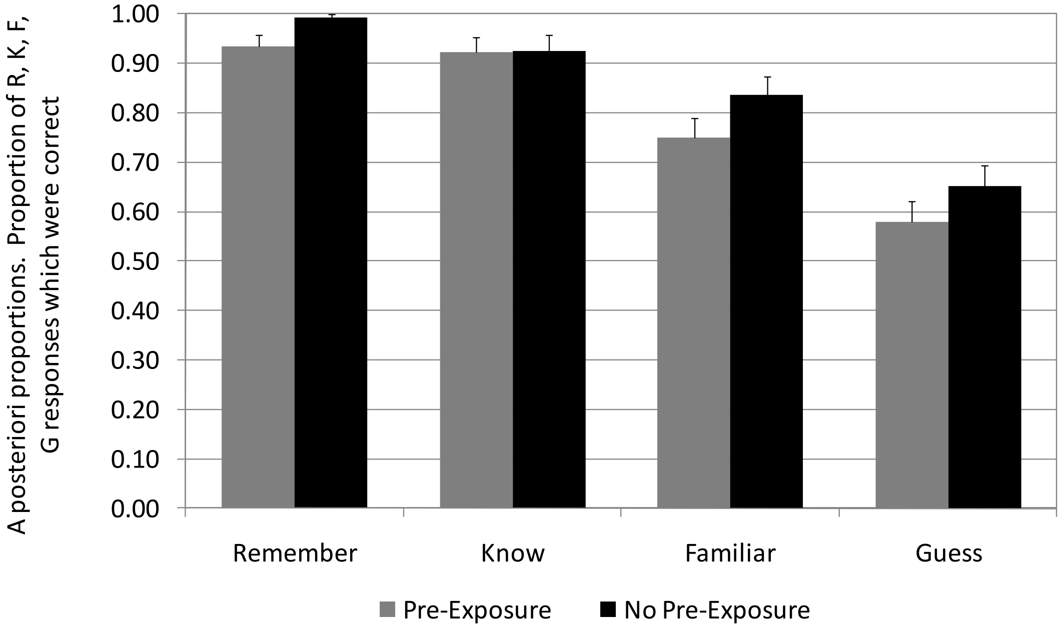

Figure 3.5. Proportion of forenames assigned to Remember, Know, Familiar, and Guess in the two conditions that had previously been correctly recognised. Errors bars = 1 SeM. N = 30 due to missing data.

A 2(exposure condition) x 4(subjective experience) within-subjects ANOVA demonstrated a significant main effect of exposure condition, $F(1,29) = 14.83$, $p = .001$, and a significant main effect of subjective experience, $F(3,87) = 43.95$, $p < .001$, but no interaction between the two, $F < 1$. Planned comparisons within conditions revealed that in the Pre-Exposure condition there was no significant difference between the likelihood of recognition having been correct when items were assigned to the Remember category compared to the Know category, $t < 1$. Conversely, in the No Pre-Exposure condition recognition decisions were significantly more likely to have been correct when items were assigned to Remember compared to Know, $t(29) = 2.06$, $p = .048$. All other comparisons between categories were significantly different at $p < .01$, except the comparison of items assigned to Know and Familiar in the No Pre-Exposure condition which was only marginally significant, $t(29) = 2.01$, $p = .054$. As was found in Experiment 3.1, the likelihood of recognition decisions being correct was always over .90 for responses which went on to be classified as Remember or Know, and for both conditions was significantly higher than the likelihood of items being correct which went on to be assigned to Familiar or Guess. Planned comparisons between the exposure conditions revealed that decisions based on Familiarity were again less accurate for Pre-Exposed items compared to items which had undergone No Pre-Exposure, and this neared conventional significance levels, $t(29) = 1.96$, $p = .06$. Additionally, in this experiment recognition decisions based on Remembering also produced significantly lower performance when items had been Pre-Exposed compared to items in the No Pre-Exposure condition, $t(29) = 2.56$, $p = .016$.

3.3.3.3. Correlational analyses

The proportion of names correctly recognised in each condition was again correlated with the assignment of those names to the different categories of subjective experience. Memory performance was found to be significantly positively correlated with proportion of items assigned to Remember in both conditions and negatively correlated with proportion of items assigned to Familiar in both conditions, see Table 3.2. Negative correlations were also found between memory performance and proportion of items assigned to Guess in both conditions. Correlations between memory performance and Know were not significant.

Table 3.2. Correlations between memory accuracy and proportion of items assigned to subjective experience categories in each condition.

Condition	Remember	Know	Familiar	Guess
Pre-Exposure	.38*	.21	-.33*	-.44**
No Pre-Exposure	.35*	.15	-.28#	-.36*

N = 44 both Pre-Exposure and No Pre-Exposure. ** $p < .005$, * $p < .03$, # $p = .06$.

To test whether there were significant differences between the correlations in the two conditions Steiger's $\bar{Z}_2$* for dependent correlations was calculated (Steiger, 1980; see Appendix D for full formulae). No significant differences were observed between the two conditions, all $\bar{Z}_2$*<1. In brief, both conditions yielded patterns of correlations similar to the Pre-Exposure condition in Experiment 3.1.

3.3.4. Discussion

Results from Experiment 3.2 replicated many aspects of those from Experiment 3.1 but they also add to them. In the current experiment pre-exposure was manipulated within-subjects by exposing participants to half the names prior to study. This manipulation was operationalised by either all or none of the three components of the 2AFC associative recognition test (target forename, studied surname, lure forename) being included on the pre-exposure vowel-counting task. Either all items were familiar – studied forename and surname seen twice, but only paired once, lure forename only seen once; or the studied forename and surname had been seen once and the lure forename not at all. Replicating Experiment 3.1, memory performance was lower for items pre-exposed prior to study compared to items which had not undergone pre-exposure. This difference demonstrates that prior familiarity with experimental items impairs memory performance using both between- and within-subjects designs.

However, in comparison to the differential use of subjective experience categories across the Pre-Exposure and No Pre-Exposure conditions in Experiment 3.1, the current results demonstrate that when pre-exposure is manipulated within-subjects, patterns of subjective experience are the same across conditions. The proportion of responses assigned to Remember, Know, Familiar, and Guess were the same regardless of whether the item had been pre-exposed or not, and the patterns paralleled those of the Pre-

Exposure condition in Experiment 3.1 with the majority of items being assigned to Remember in both conditions. This same pattern of usage of subjective experience categories in the two conditions was also reflected in the correlational analysis which showed that use of Remembering as a basis for recognition judgments was positively related to correct recognition for both pre-exposed and not pre-exposed names, which differs from Experiment 3.1 where this was only found in the Pre-Exposure condition.

Analysis of *a posteriori* proportions also demonstrated differences from the results found in Experiment 3.1. In the current experiment, when participants made their recognition decisions on either the basis of Remembering, or on the basis of Familiarity, both these decisions were more likely to be incorrect if the item had been on the pre-exposure list. The finding regarding familiarity-based recognition decisions parallels that found in Experiment 3.1; however, in that experiment no differences were found for recollection-based decisions. Therefore, using both between- and within-subjects designs, when items are pre-exposed, basing a recognition decision on familiarity was found to significantly impair performance, however, only in the within-subjects design was this also true for recollection-based decisions. Here, when some target and lure forenames had all been pre-exposed but others had not, recollection of a lure forename from its presence in the pre-exposure task may have led to it being selected over the target forename.

In this experiment, correlational analysis demonstrated memory performance to be positively related to proportion of items assigned to Remember in both conditions and negatively related to proportion of items assigned to Familiar in both conditions. This suggests that for all items in this within-subjects design, pre-exposed or not, it was necessary to use recollection to resolve the familiarity induced by pre-exposure. This initially seems to contradict the *a posteriori* data above, which indicates that recollection can impair recognition when items have been pre-exposed. However, taken together these results suggest that overall recollection aids recognition performance for all items when pre-exposure is manipulated within-subjects, but for a small proportion of items when the manipulation of pre-exposure is within-subjects incorrect recollection of pre-exposed lures can impair recognition in the Pre-Exposure condition.

In summary, the memory performance advantage for items not pre-exposed replicated findings from Experiment 3.1 and findings from previous experiments on word frequency and experimental manipulations of familiarity. However, results regarding use of subjective experience categories did not match those obtained in Experiment 3.1 and it is suggested that this is due to carry-over effects from the Pre-Exposure condition to the No Pre-Exposure condition in this within-subjects design. As discussed by Conway et al. (2001) one problem in within-subjects designs is the potential for contamination of one condition by the other condition. Indeed, this is exactly what was observed in the Remember-Know experiments performed by Conway et al. (2001). In the current experiment, it appears that manipulating pre-exposure using a within-subjects design led participants to use the same underlying recognition processes or subjective experiences to make *all* their recognition decisions, regardless of whether items had been pre-exposed or not. The difference in the relative familiarity of target and lure names across pre-exposed and non-pre-exposed items was too subtle to be used by participants to judge for which items recollection was necessary for recognition and for which items it was not. The patterns observed in both conditions of Experiment 3.2 matched those from the Pre-Exposure condition of the between-subjects Experiment 3.1 suggesting that the patterns of subjective experience use 'carried-over' from the Pre-Exposure condition to the No Pre-Exposure condition. This also relates to the patterns demonstrated by Whittlesea (2002). He found that subjective experience responses were dependent not only on the familiarity of the item itself, but also on how the item was perceived in relation to other items encountered at test. Results from Experiments 3.1 and 3.2 demonstrate that the different contexts in which pre-exposed and non-pre-exposed items are encountered influences how those items are subjectively experienced. These issues will be returned to in the General Discussion (Sections 3.5.2 and 3.5.4).

3.4. Experiment 3.3: Recollection in item and non-associative recognition in a blocked within-subjects design

3.4.1. Introduction

This experiment aimed to extend the findings of Experiments 3.1 and 3.2 in a number of ways by using a blocked within-subjects design. By manipulating pre-exposure using a blocked design this experiment aimed to examine whether the carry-over of use of subjective experience from pre-exposed to non-pre-exposed items would occur when participants experienced both types of items but they were separated across different blocks.

It was suggested that in Experiment 3.2, when pre-exposure was manipulated within the one experimental block, participants were not able to tell which items had been pre-exposed and consequently their use of subjective experiences to judge recognition was the same for both types of item. In the current experiment, if participants are able to tell that in one block the lure words are completely novel whereas in another block the lure words are words they just performed the pre-exposure task on, then their patterns of subjective experience should differ across blocks. In other words, if participants are aware of the differences in relative familiarity of items across blocks then this should be reflected in their judgments of subjective experience. Conversely, if participants are unable to perceive the differences across blocks then patterns of subjective experience should show evidence of carry-over from pre-exposed to non-pre-exposed blocks. As discussed previously (Section 3.1.5), Bodner and Richardson-Champion (2007) recently explored carry-over effects and test-list context in an experiment on eyewitness testimony. They found that the relative difficulty of items on the test influenced subjective experience responses when difficulty was manipulated within list. In addition, when detail difficulty was manipulated across blocks the level of difficulty of one block influenced the subjective experience responses given to items in the next block. In the current experiment, if carry-over effects are observed then these are able to be examined by exploring the order in which participants completed the different blocks.

Two further aims of this experiment were to see whether the previous findings would be replicated with words rather than names as stimuli and whether the findings obtained using an associative-2AFC paradigm would be replicated in non-associative-2AFC recognition and item (Yes/No) recognition. To these ends four experimental blocks were tested. The first two blocks employed a Yes/No recognition test and the latter two blocks employed a non-associative-2AFC test; full description of the design is given in Section 3.4.2.2.

Previous studies have found mixed results regarding whether use of recollection is greater in item versus AFC recognition. In a test of face recognition, Bastin and Van der Linden (2003) found recollection to be greater in item recognition compared to AFC recognition.

This fits with the widespread assumption that AFC recognition is thought to rely less on recollection than item recognition (e.g., Aggleton & Shaw, 1996) as in AFC recognition one can use the relative familiarity of items to make a recognition decision (e.g., Glanzer & Adams, 1990; Glanzer & Bowles, 1976). However, Khoe, Kroll, Yonelinas, Dobbins, and Knight (2000) and Kroll, Yonelinas, Dobbins, and Frederick (2002) found no differences in the amount of recollection in AFC and item recognition. Interestingly from the point of the current experiments, Cook et al. (2005) found contributions of recollection to AFC and item recognition to differ according to word frequency. Using a between-subjects design they found that for low-frequency words more items were assigned to Remember in item compared to AFC recognition, whereas for high-frequency words more were assigned to Remember in AFC than in item recognition. In the current experiment, patterns of subjective experience responses were compared across item and 2AFC recognition to explore whether pre-exposure influenced subjective experiences differentially in these two tasks.

3.4.2. Method

3.4.2.1. Participants

Participants were 37 students and staff from the University of Leeds (31 female) with a mean age of 20.6 years who were members the Institute Participant Pool or on an email list of people willing to take part in Psychology experiments. They received either Participant Pool Credits or £6 payment for taking part. As an additional incentive to concentrate and perform as well as possible across this rather long experiment, participants received a further £5 if they achieved over 85% memory accuracy overall. Fifteen participants achieved this and were compensated. Participants were tested in groups of between three and nine at individual PCs in a laboratory computer cluster. All instructions and stimuli were presented, and data collected, using E-Prime version 1.2.

3.4.2.2. Design and Materials

The experiment employed a blocked within-subjects design. The experiment consisted of four blocks, each block consisting of a pre-exposure phase, study phase and test phase. Critically, in only two of the four blocks was the pre-exposure performed on the words that were then to be the target and lure words, in the other two blocks the ‘pre-exposure’ task was performed on distinct, i.e., non-studied, word lists. These blocks were the No Pre-

Exposure blocks. For all participants the first two blocks (one Pre-exposure, one No Pre-Exposure) were tested using Yes/No recognition tests whereas blocks three and four employed 2AFC recognition tests (again one Pre-exposure, one No Pre-Exposure). In contrast to the design employed in Experiments 3.1 and 3.2 the 2AFC task here was not an associative recognition test. In the previous experiments participants studied two items together, a forename and a surname, and were presented with the surname and two forenames at test from which they had to choose the forename they had studied with that surname. Here items for study were presented individually and in the 2AFC test the correct response was presented together with a lure item.

The stimuli were 480 medium- to high-frequency words[13] (mean familiarity rating of 545; range 488-652) limited to between four and seven letters in length. The words were pseudo-randomly assigned to list to achieve approximately equal alphabetical spread across 12 lists of 40 words each (word lists shown in Appendix B.3). Eight versions of the experiment were created by counterbalancing which lists were to be used as target and lure words, and in which order participants would perform the Pre-Exposure blocks and the No Pre-Exposure blocks. Full explanation of this counterbalancing is shown in Appendix B.3. Pairing of target and lure words for the 2AFC blocks was also pseudo-random with matching initial letters avoided. In these blocks the side of the screen on which the targets and lures appeared was also counterbalanced. Presentation of all stimuli in all phases was randomised by E-Prime.

3.4.2.3. Procedure

Participants were instructed about the general procedure of the experiment and informed consent was obtained. Participants then accessed the experiment file and all further instructions for each phase were presented via computer screen. Participants completed four experimental blocks each consisting of a pre-exposure, study, and test phase.

For the pre-exposure phases, participants were instructed that a word would be shown in the centre of the screen with a blue line below it and that their task was to count the

[13] Words obtained from the MRC Psycholinguistic Database. Familiarity values refer to the printed frequency in the language and were derived from merging three sets of familiarity norms: Pavio (unpublished), Toglia and Battig (1978) and Gilhooly and Logie (1980).

number of vowels in the word. Participants were instructed to use the number keys to input their response and that it would appear on the blue line as they typed. To proceed to the next word participants then pressed the enter key. For all trials the instruction "Count the number of vowels in this word" appeared at the top of the screen and "Press enter to continue" appeared at the bottom. In order that participants completed each trial quickly, a time limit of four seconds was given for each trial. If the participant took longer than four seconds to make a response the word disappeared and a message saying "Too Slow" was shown on the screen before the computer moved onto the next trial. In each pre-exposure phase, participants completed this vowel-counting task on 80 words (40 targets plus 40 lures in the Pre-Exposure conditions, and 80 unrelated words in the No Pre-Exposure conditions).

In the study phase, the 40 target items were presented individually on the screen for four seconds with a fixation point shown for 750 milliseconds between items. Participants were instructed that they should study the words and that later their memory would be tested. For each block participants were instructed, prior to study, which type of recognition test they would undergo in the test phase of that block: a Yes/No recognition test or a 2AFC recognition test.

In each block, when the study phase was completed participants moved directly on to the test phase. Participants were again instructed as to which type of test was involved in this block, 2AFC or Yes/No and examples of the screen layouts were shown. Participants were also given time to familiarise themselves with the definitions of Remember (R), Know (K), Familiar (F), and Guess (G) which they were to use to judge their recognition experience for all items (see Appendix A.3). Each experimental trial began with a fixation point being shown for 750ms. In the blocks that had a Yes/No recognition test, a word was then presented in the centre of the screen with boxes labelled 'Yes' and 'No' positioned below it and towards the left and the right respectively. The instruction "Did you study this word in the Study Phase? Use the mouse to respond" appeared at the top of the screen and participants clicked on one of the boxes to make their response. The mouse pointer returned to the centre of the screen at the start of each trial. The 40 target words and 40 lures were presented in random order. In the test phases of the 2AFC blocks, a target and a lure word were presented to the left and right of the centre of the screen (placement counterbalanced) with the instruction "Which of these words did you study in the Study

Phase? Use the mouse to respond" and participants clicked on the word they had studied previously. Again the mouse pointer returned to the centre of the screen at the start of each trial. After making their recognition decision, participants were shown the subjective experience judgment screen which consisted of four response boxes labelled R, K, F, and G, with the question "Please make a recognition rating" and made a subjective experience judgment by clicking the mouse on one of the boxes. Speed and accuracy of responses were emphasized in the instructions for the test phase and response times for recognition decisions and subjective experience judgments were recorded. Prior to a new trial commencing a blank screen was shown for 750ms.

In between blocks participants were instructed that it was the end of the block and told "You will not be asked to remember any of those words again, the next block involves a completely new set of words." They were also permitted to take a moment to take a break if they were feeling tired, and they began the next block when ready. After completion of the four blocks participants were debriefed.

3.4.3. Results

To determine whether pre-exposure or task type influenced memory performance recognition performance in each of the four blocks was examined. Within each block, proportions of correct responses assigned to each of the subjective experience categories (Remember, Know, Familiar, and Guess) were again calculated using both *a priori* and *a posteriori* methods. To examine whether the order in which initial Yes/No Pre-Exposure and No Pre-Exposure blocks were performed had influenced data in these conditions, memory performance and use of subjective experience categories were also analysed split by counterbalancing group. Correlational analyses were also carried out. Finally, recognition decision reaction times were analysed by their later subjective experience category. Prior to analysis, data from any responses which had been made faster than 300ms or slower than 8000ms were excluded from the dataset.

3.4.3.1. Memory measures

The proportion of items correctly recognised in the Pre-Exposure and No Pre-Exposure conditions in the Yes/No task and the 2AFC task were examined. As can be seen from Table 3.3, in both tasks recognition performance was lower when items had been pre-exposed.

As expected, raw proportions of recognition performance were much lower in the Yes/No task as opposed to the 2AFC task, however this difference was eliminated when performance was compared using *d'* as the formula for calculating *d'* in 2AFC recognition takes into account the performance advantage of 2AFC over Yes/No recognition[14].

Table 3.3. Means (and standard deviations) of recognition performance and *d'* values in each of the blocks[14].

	Yes/No Task		2AFC Task	
Condition	Recognition performance	*d'* value	Recognition performance	*d'* value
Pre-Exposure	.61 (.23)	1.91 (.97)	.89 (.14)	1.94 (.88)
No Pre-Exposure	.71 (.20)	2.36 (.98)	.91 (.11)	2.06 (.78)

A 2(task) x 2(exposure condition) within-subjects ANOVA on recognition performance demonstrated a main effect of task, $F(1,36) = 120.19$, $p < .001$, confirming that when looking at raw proportions memory performance was significantly lower in the Yes/No task compared to the 2AFC task; a main effect of exposure condition, $F(1,36) = 13.44$, $p = .001$; and an interaction between exposure and task, $F(1,36) = 5.66$, $p = .02$. Planned comparisons within task revealed that the effect of exposure condition was mainly due to the significant difference in memory performance in exposure conditions in the Yes/No task, $t(36) = 3.31$, $p = .001$ (one-tailed), as the difference in the means in the 2AFC task only approached significance, $t(36) = 1.62$, $p = .057$ (one-tailed). Although mean recognition performance in the 2AFC task was high, one-sample t-tests against the maximum score of 1 demonstrated performance in both conditions was significantly below ceiling, both $p < .001$.

The corresponding 2(task) x 2(exposure condition) within-subjects ANOVA on the *d'* values showed no main effect of task, $F(1,36) = 1.84$, $p = .18$, demonstrating that the adjustment used when calculating *d'* in 2AFC did eliminate the performance advantage typically seen for 2AFC recognition over Yes/No recognition. ANOVA did show a main effect of exposure condition, $F(1,36) = 12.23$, $p = .001$ and an interaction between exposure and task, $F(1,36)$

[14] In the Yes/No blocks the calculation used for recognition performance was: (hits-false alarms)/number of items; whereas in the 2AFC blocks it was = hits/number of items. The different formulas for *d'* in Yes/No and 2AFC recognition are shown in Appendix C.

= 4.05, p = .05. To avoid repetition further comparisons were not performed using d'. As differences between Yes/No and 2AFC recognition were not the primary interest in this design, further analysis does not compare tasks but was performed on the Yes/No and 2AFC data separately.

3.4.3.2. Exposure condition order effects on memory performance

To examine whether there were any carry-over effects seen in this blocked design whereby performance in participants' second block was influenced by which condition they had experienced in their first block, memory performance on the Yes/No task was examined with participants' counterbalancing group as a between-subjects factor. The means in Table 3.4 show that memory performance in the Pre-Exposure condition was approximately equal (around .60) whether that condition was performed as the first or second block. For No Pre-Exposure, mean performance was higher, .76 compared to .65, if this block was performed second as opposed to first. A 2(exposure condition) x 2(block order) ANOVA revealed no between-subjects main effect of block order, $F < 1$, but a significant main effect of exposure, $F(1,35) = 11.97$, $p = .001$, and a significant interaction between exposure and block order, $F(1,35) = 6.02$, $p = .02$.

Table 3.4. Means (and standard deviations) of recognition performance in Pre-Exposure and No Pre-Exposure conditions split by which exposure condition was performed first or second. N = 19 for participants who performed Pre-Exposure first. N = 18 for participants who performed Pre-Exposure second.

Condition	1st Block	2nd Block
Pre-Exposure	.59 (.23)	.62 (.24)
No Pre-Exposure	.65 (.18)	.76 (.21)

Firstly a t-test was performed on only the first block data, as this comparison would be directly comparable to the between-subjects manipulation of pre-exposure in Experiment 3.1. The difference in mean memory performance here (means .59 and .65 for Pre-Exposure and No Pre-Exposure respectively), was not found to be significant, $t(35) = .97$, $p = .17$ (one-tailed). This is probably largely due to the reduction in N from splitting participants by their counterbalancing group. Comparisons within condition demonstrated no significant differences for performance dependent on whether block was carried out first or second for Pre-Exposure, $t(35) = .48$, $p = .63$, or No Pre-Exposure, $t(35) = 1.63$, $p = .11$. Contrastingly, comparisons within block order revealed memory performance was significantly better in the No Pre-Exposure condition if this condition was performed second, after the Pre-Exposure condition (.76 compared to .59), $t(18) = 4.28$, $p < .001$. However, for participants who underwent Pre-Exposure second after No Pre-Exposure there was no significant difference in memory performance (.62 compared to .65), $t(17) = .70$, $p = .50$. This demonstrates that the order in which participants performed the blocks influenced their memory performance. For participants who experienced pre-exposure first, their memory performance increased significantly in the non-pre-exposure block that followed. In contrast, for participants who experienced the non-pre-exposure block first, their memory performance did not differ when they then encountered the pre-exposure block.

3.4.3.3. Subjective experience measures at retrieval – a priori proportions

The next analysis performed was to examine what type of subjective experience participants had been basing their recognition decisions on. *A priori* proportions were calculated – the proportion of correct responses assigned to each of the subjective experience categories. As is shown in Figure 3.6, correct recognition was based on similar usage of subjective experience across both exposure conditions and tasks. In all blocks, participants assigned nearly 50% of correctly recognised names to the Remember category, approximately 32% to Know, 15% to Familiar, and less than 6% to Guess.

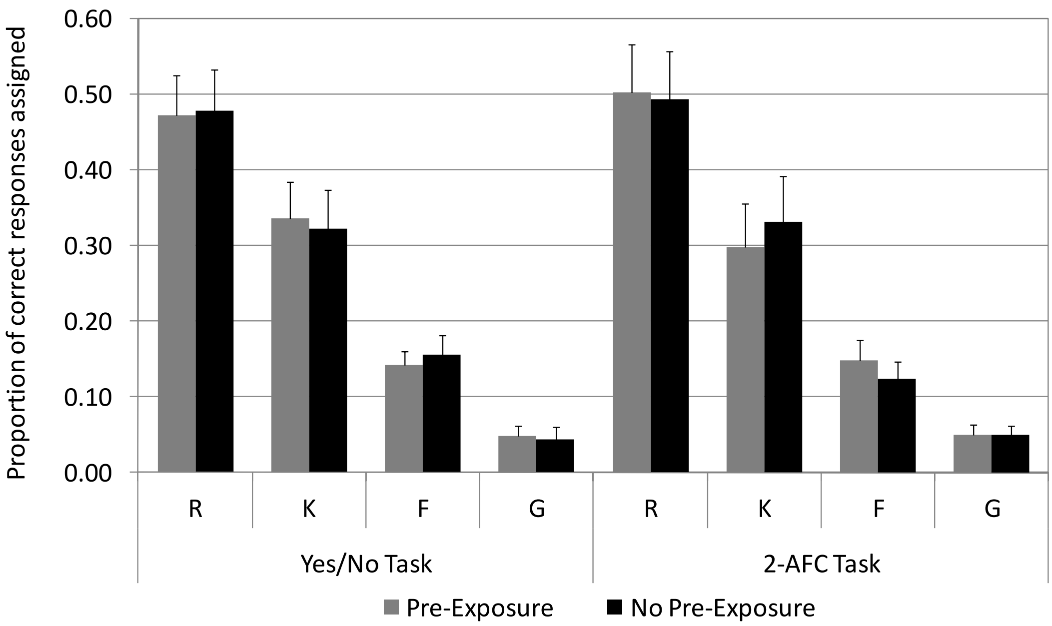

Figure 3.6. Proportion of correct responses assigned to Remember (R), Know (K), Familiar (F), and Guess (G) by participants in the Pre-Exposure and No Pre-Exposure conditions of the Yes/No task and the 2AFC task. Errors bars = 1 SeM.

Separate 2(exposure condition) x 4(subjective experience) within-subjects ANOVAs within each task demonstrated a significant main effect of subjective experience in both the Yes/No task, $F(3,108) = 19.34$, $p < .001$, and the 2AFC task, $F(3,108) = 14.92$, $p < .001$. Effects of exposure condition could not be calculated as proportions summed to 1 within conditions. There were no significant interactions between exposure condition and subjective experience, both $F < 1$.

Planned comparisons showed there were no significant differences between the proportion of responses assigned to Remember, Know, Familiar, and Guess in the different exposure conditions, all $p > .20$. Comparisons within conditions revealed no significant difference between the amount of responses assigned to Remember compared to Know in: Yes/No Pre-Exposure, $t(36) = 1.38$, $p = .18$; Yes/No No Pre-Exposure, $t(36) = 1.55$, $p = .13$; and 2AFC No Pre-Exposure, $t(36) = 1.34$, $p = .19$. However the difference between proportion assigned to Remember and Know in the 2AFC Pre-Exposure condition approached significance, $t(36) = 1.77$, $p = .09$. Although the means for proportion of responses assigned to Remember in Figure 3.6 appear to be much higher than those for Know, the large standard errors mean that these differences did not reach significance; participants assigned statistically identical proportions of responses to Remember and Know in all conditions. Comparisons of all other differences were significant, all at least $p < .02$.

3.4.3.4. Exposure condition order effects on subjective experience at retrieval

As memory performance was shown to interact with which order participants experienced conditions (see Section 3.4.3.2), the influence of block order on subjective experience at retrieval was also explored. Proportion of correct responses assigned to Remember, Know, Familiar, and Guess in the Pre-Exposure and No Pre-Exposure conditions of the Yes/No task were split by which order they performed the conditions in; means shown in Figure 3.7.

A 2(exposure condition) x 2(block order) x 4(subjective experience) ANOVA demonstrated a significant main effect of subjective experience, $F(3,105) = 19.13$, $p < .001$. Effects of exposure and block order could not be computed as proportions summed to 1; however there was a significant three-way interaction between exposure, block order, and subjective experience, $F(3,105) = 3.74$, $p = .013$. Examining the means in Figure 3.7 this interaction would appear to be based on the differences in proportion of responses assigned to Remember, and perhaps also Familiarity, in the different exposure conditions dependent on which order the conditions were undertaken. A greater proportion of items were assigned to Remember in the No Pre-Exposure condition when it followed the Pre-Exposure condition (left panel), whereas if Pre-Exposure followed No Pre-Exposure the patterns are reversed and a greater proportion of items were assigned to Remember in Pre-Exposure (right panel).

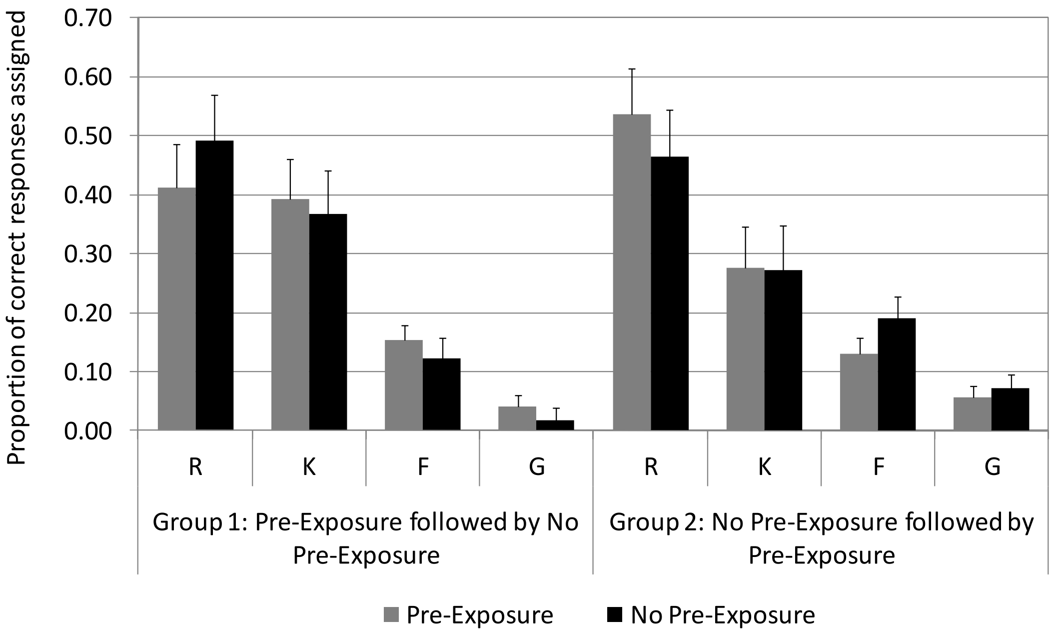

Figure 3.7. Proportion of correct responses assigned to Remember (R), Know (K), Familiar (F), and Guess (G) by participants in the Pre-Exposure and No Pre-Exposure conditions of the Yes/No task split by which order they performed the conditions in. N = 19 for participants who performed Pre-Exposure first. N = 18 for participants who performed Pre-Exposure second. Errors bars = 1 SeM.

Separate 2(exposure condition) x 2(block order) ANOVAs were performed for each of the subjective experience categories. For both Remember and Familiar, there were no main effects of exposure or block order, however significant interactions between block order and exposure were found, Remember: $F(1,35) = 5.52$, $p = .025$, and Familiar: $F(1,35) = 4.96$, $p = .032$. This suggests that the counterbalancing of which order participants performed the Pre-Exposure or No Pre-Exposure blocks in influenced the subjective experience used to make recognition decisions. Participants used different strategies depending on which block they experienced first. The influence of counterbalancing on subjective experience and memory performance is returned to in the Discussion, Section 3.5.4.

3.4.3.5. Subjective experience measures at retrieval – a posteriori proportions

To explore whether, in any condition, an item that had been assigned to a particular subjective experience category was more likely to have come from a correct or incorrect recognition response *a posteriori* probabilities were also calculated for this experiment[15]. Due to high performance, particularly in the 2AFC task, a lot of missing data was encountered as participants had not assigned responses to all subjective experience categories. As shown in Section 3.4.3.3, very few responses were assigned to Guess and therefore to increase N this category was excluded from analysis.

As can be seen in Figure 3.8, recognition was more likely to have been accurate when it was later assigned to Remember or Know. In addition, for the majority of categories the probability of an item coming from a correct recognition response was lower for items in the Pre-Exposure condition than in the No Pre-Exposure condition. Both these patterns replicate those observed in Experiments 3.1 and 3.2. Comparing Yes/No and 2AFC probabilities it is evident that accuracy of items assigned to Remember and Know was higher in the 2AFC task (where overall memory performance was near ceiling) compared to in the Yes/No task.

Separate 2(exposure condition) x 3(subjective experience) within-subjects ANOVAs were performed for the Yes/No task and the 2AFC task. ANOVA for the Yes/No task demonstrated a significant main effect of exposure condition, $F(1,23) = 5.06$, $p = .03$, a significant main effect of subjective experience category, $F(2,46) = 17.30$, $p < .001$, and an interaction between the two, $F(2,46) = 3.25$, $p < .05$.

Planned comparisons within condition for the Yes/No task showed that in both the Pre-Exposure and No Pre-Exposure conditions there was no significant difference between the likelihood of recognition decisions having been correct when items were assigned to the Remember category compared to the Know category, both $t < 1$, whereas all other comparisons between categories were significant, all at least $p < .02$. The likelihood of recognition decisions being correct was high (over .75 in Pre-Exposure and over .85 in No

[15] In light of the effects of counterbalancing observed in the *a priori* analysis, this *a posteriori* analysis was also conducted using a 2(exposure) x 2(block order) x 3(subjective experience category) ANOVA however no 3-way interaction was observed, $F < 1$, so block order was removed as a factor.

Pre-Exposure) for responses which went on to be classified as Remember or Know, and for both conditions was significantly higher than the likelihood of items being correct that were assigned to Familiar. Comparisons between the conditions revealed that for recognition decisions that were assigned to Remember or to Know, responses in the No Pre-Exposure condition were significantly more likely to have been correct than responses in the Pre-Exposure condition, Remember: $t(23) = 3.91$, $p = .001$, Know: $t(23) = 3.00$, $p = .006$. When participants made their recognition decisions on the basis of Familiarity the difference between conditions was not significant, $t < 1$.

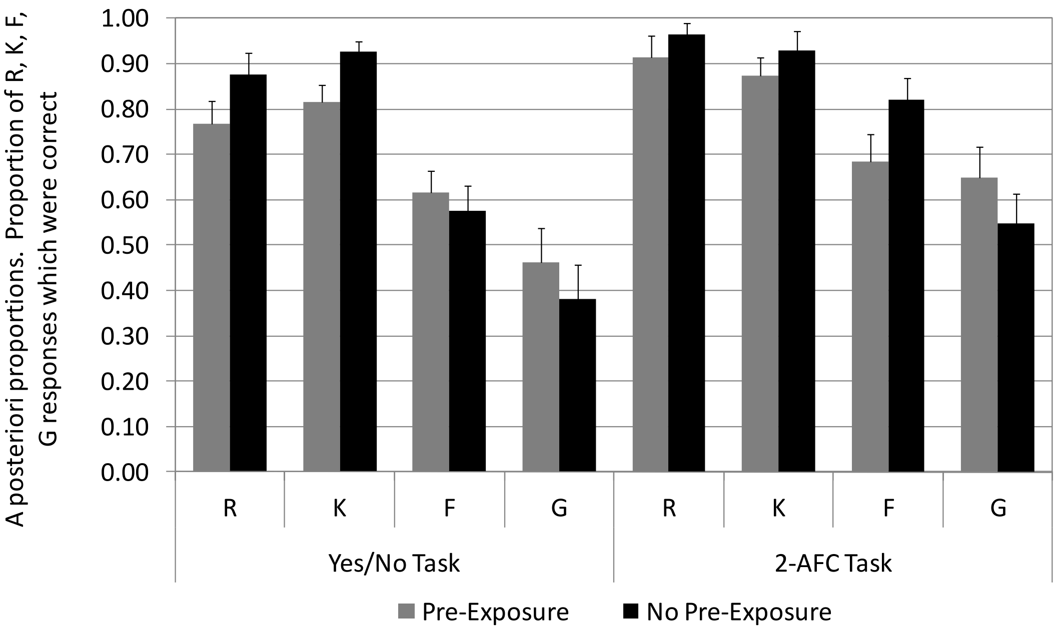

Figure 3.8. Proportion of items assigned to Remember (R), Know (K), Familiar (F), and Guess (G) in the in the Pre-Exposure and No Pre-Exposure conditions of the Yes/No task and the 2AFC task that had previously been correctly recognised. For R, K, and F: N = 24 for Yes/No and N = 14 for 2AFC due to missing data[16]. Proportions for Guess shown but not included in ANOVA. Errors bars = 1 SeM.

The 2(exposure condition) x 3(subjective experience) within-subjects ANOVA for the 2AFC task revealed a significant main effect of exposure condition, $F(1,13) = 8.13$, $p = .01$, and a significant main effect of subjective experience category, $F(2,26) = 10.26$, $p = .001$, but no interaction between the two, $F(2,26) = 1.09$, $p = .35$. Planned comparisons within condition

[16]Although N is greatly reduced by listwise exclusion for purposes of ANOVA, the inclusive means show the same patterns and are shown in Appendix H.

for the 2AFC task showed that in both the Pre-Exposure and No Pre-Exposure conditions there was no significant difference between the likelihood of recognition decisions having been correct when items were assigned to Remember compared to Know, both $t < 1$, whereas all other comparisons between categories were significant, all at least $p < .02$, except the comparison of Know and Familiar in No Pre-Exposure which was marginally significant, $t(13) = 2.01$, $p = .07$. The likelihood of recognition decisions being correct was consistently high, over .87, for responses which went on to be classified as Remember or Know, and again for both conditions was higher than the likelihood of items being correct which went on to be assigned to Familiar. Comparisons between the conditions revealed contrasting results to those in the Yes/No task. Here it was only for recognition decisions which were assigned to Familiar that responses in the No Pre-Exposure condition were significantly more likely to have been to a correct response than responses in the Pre-Exposure condition, $t(13) = 2.72$, $p = .02$. When participants made their recognition decisions on the basis of Remember or Know subjective experiences the difference between conditions was not significant, Remember: $t(13) = 1.14$, $p = .28$, Know: $t(13) = 1.21$, $p = .25$.

3.4.3.6. Correlational analyses

Again, to examine whether recognition based on a particular type of subjective experience was related to performance, memory accuracy in each condition was correlated with proportion of correct responses assigned to the different levels of subjective experience, see Table 3.5.

Table 3.5. Correlations between memory accuracy and proportion of items assigned to subjective experience categories in each condition.

Task	Condition	Remember	Know	Familiar	Guess
Yes/No task	Pre-Exposure	.34*	.00	-.57**	-.53**
	No Pre-Exposure	.60**	-.16	-.67**	-.45**
2AFC Task	Pre-Exposure	.35*	.03	-.62**	-.49**
	No Pre-Exposure	.30#	.13	-.82**	-.67**

df = 37 in all conditions. ** $p < .005$, * $p < .04$, # $p = .07$.

Across both tasks and both exposure conditions memory performance was found to be positively correlated with proportion of correct items assigned to Remember, although in the No Pre-Exposure condition of the 2AFC task this correlation only approached significance, $r(37) = .30$, $p = .07$; all others at least $p < .05$. In all conditions memory performance was also negatively correlated to proportion of items assigned to Familiar and Guess, all $p < .001$. Proportion of items assigned to Know never significantly correlated to memory performance[17].

To test whether there were significant differences between the correlations Steiger's $\bar{Z}_2$* (Steiger, 1980) was again calculated (see Appendix D). Although this does not allow for full exploration of the interactions between all coefficients, it does allow us to see whether key comparisons are significantly different. Comparing the coefficients for the subjective experience categories across the two conditions revealed a significant difference between the coefficients for Remember in the Yes/No task. The positive correlation between memory performance and proportion of items assigned to Remember was significantly larger in the No Pre-Exposure condition than the Pre-Exposure condition, $\bar{Z}_2* = 1.97$, $p < .05$. No other coefficients were significantly different in this task, all $\bar{Z}_2* < 1.1$. However, in the 2AFC task the negative correlation between memory performance and proportion of items assigned to Familiar was significantly larger in the No Pre-Exposure condition than the Pre-Exposure condition, $\bar{Z}_2* = 2.47$, $p < .05$. No other coefficients were significantly different, all $\bar{Z}_2* < 1.6$.

3.4.3.7. Reaction times for recognition decisions

As this experiment was conducted using E-Prime, RTs to make recognition decisions categorised by their later subjective experience judgment could be analysed as per Experiment 3.1. As discussed previously, very few responses were assigned to Guess and therefore to increase N this category was excluded from analysis. As is shown in Figure 3.9, mean RTs for correct recognition decisions were faster for responses later categorised as Remember or Know than those categorised as Familiar or Guess which replicates the patterns from Experiment 3.1.

[17] Due to the loss of power associated with splitting this data by block order, analysis of counterbalancing effects was not performed.

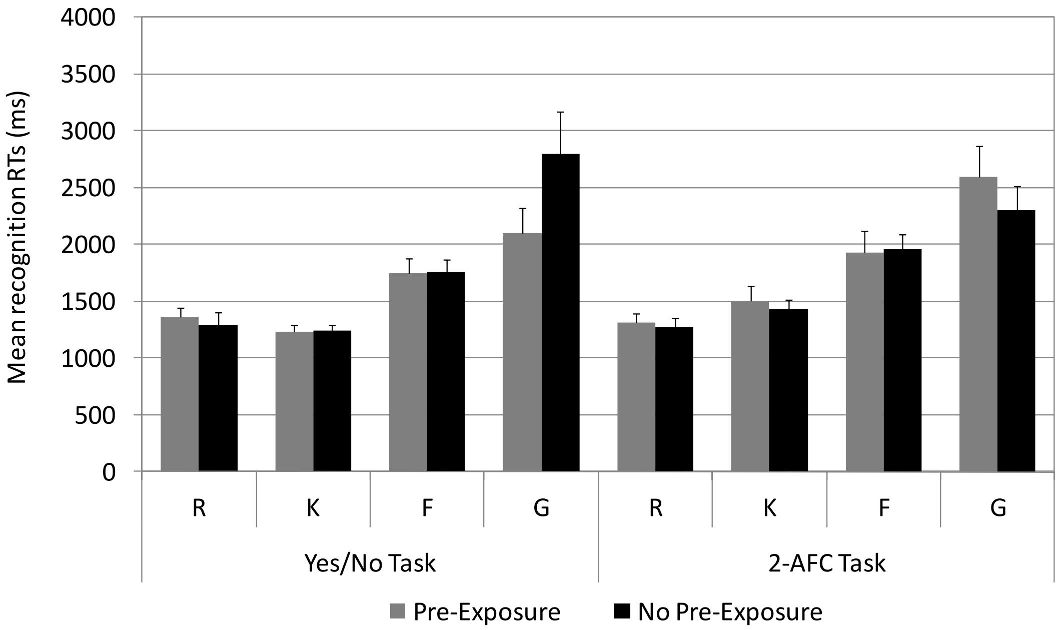

Figure 3.9. Mean correct recognition RTs by later subjective experience category, Remember (R), Know (K), Familiar (F), and Guess (G), in the in the Pre-Exposure and No Pre-Exposure conditions of the Yes/No task and the 2AFC task. Error bars = 1 SeM. For R, K, and F: N = 22 for Yes/No and N = 14 for 2AFC due to missing data[18]. Proportions for Guess shown but not included in ANOVA.

Separate 2(exposure condition) x 3(subjective experience) within-subjects ANOVAs were performed for the Yes/No task and the 2AFC task. In both tasks significant effects of subjective experience category were observed, Yes/No: $F(2,42) = 17.83$, $p < .001$; 2AFC: $F(2,26) = 18.61$, $p < .001$. No effects of exposure condition or interactions between exposure and subjective experience were revealed in either task, all $F < 1$.

Collapsed across exposure condition, comparisons demonstrated that in the Yes/No task the speed with which recognition decisions were made which were later classed as Remember or Know was not significantly different, $t(21) = .97$, $p = .34$, however Remember and Know decisions were significantly faster than Familiar decisions, both $p < .001$. In the 2AFC task however, recognition decisions were made significantly faster if they were later classed as Remember rather than Know, $t(13) = .2.53$, $p = .03$, and again Remember and Know decisions were significantly faster than Familiar decisions, both $p < .001$.

[18]Although N is greatly reduced by listwise exclusion for purposes of ANOVA, the inclusive means show the same patterns and are shown in Appendix H.

3.4.4. Discussion

3.4.4.1. Memory performance

In the current experiment, manipulation of pre-exposure was carried out using a blocked within-subjects design and both Yes/No and 2AFC recognition. Using this design, prior exposure was demonstrated to impair memory performance in the Yes/No recognition task but not significantly in the 2AFC task. This latter finding contradicts the findings from Experiments 3.1 and 3.2, however, the 2AFC design in this experiment was subtly different from the prior experiments. Experiments 3.1 and 3.2 had used associative recognition using names as stimuli whereas in Experiment 3.3 the 2AFC task was non-associative and words were employed as stimuli. These differences appear to have made the task easier for participants as performance increased in the 2AFC task in this experiment to near ceiling, eliminating significant evidence of the exposure effect.

Analysis regarding counterbalancing order effects on memory performance were carried out to determine whether the order in which participants had performed the initial two Yes/No blocks had influenced memory performance in these blocks. Firstly comparisons were made on only the first block data, as this would be directly comparable to the between-subjects manipulation of pre-exposure in Experiment 3.1. Although the difference in mean memory performance was considerable, it was not found to be significant, though it is suggested that this is due to loss of power associated with splitting participants by their counterbalancing group and conducting a between-subjects analysis (Conway et al., 2001).

Comparisons within block order however revealed memory performance was significantly better in the No Pre-Exposure condition if this condition was performed second, after the Pre-Exposure condition. This suggests that if the first block encountered by the participant is Pre-Exposure, the fact that they have to overcome the experimenter-introduced familiarity in this block leads to improvement in performance in the second block where there is no familiarity that needs to be overcome. This influence of one condition on another is discussed in detail later in relation to analysis of use of subjective experience categories split by counterbalancing block.

For participants who performed the Pre-Exposure block second, after the No Pre-Exposure block there was no significant difference in memory performance across blocks. This is somewhat surprising as it might be expected that the standard advantage in memory performance in No Pre-Exposure would be shown if this is the first block, and that this would be followed by a reduction in performance when Pre-Exposure is performed in the second block, as this is where the confusion of where an item is familiar from has to be resolved for the first time. It is suggested that one reason this was not found here is that although items in the No Pre-Exposure condition were not themselves pre-exposed, participants were exposed to a distracter set of words prior to study in this condition. This 'irrelevant' pre-exposure could have led to some problems of source discriminability being encountered in this condition, leading to a reduction in performance. A follow-up experiment could explore this further by comparing irrelevant pre-exposure (distractor words), relevant pre-exposure (lure and target words pre-exposed), and no pre-exposure in a blocked design. In this experiment, when a 'pure' no pre-exposure block is experienced first it would be expected that memory performance would be high and would be followed by a reduction in performance on subsequent blocks when any kind of pre-exposure (irrelevant or relevant) is encountered.

3.4.4.2. Subjective experience

In line with the results from Experiment 3.2, in the current experiment where pre-exposure was manipulated using a blocked within-subjects design, subjective experience was not found to differ according to exposure condition. *A priori* use of recollection was found to be the same in both conditions even though memory performance was lower with pre-exposure. Even using this blocked design, it is evident that use of subjective experience as a basis for recognition was carried-over from the Pre-Exposure condition to the No Pre-Exposure condition. Further evidence of this was obtained from the analysis split by which order participants had completed the different exposure blocks.

Analysis of use of subjective experience in the Yes/No blocks split by counterbalancing group suggested that the fact that block order influenced memory performance could be due to differing processes being used to make recognition decisions by the different groups. A greater proportion of items were found to be assigned to Remember, and less to Familiar in the No Pre-Exposure condition when it followed the Pre-Exposure condition. However, if Pre-Exposure followed No Pre-Exposure the patterns were reversed and a greater proportion of items were assigned to Remember, and less to Familiar, in the Pre-Exposure condition. These results fit with the interpretations of the effect of block order on memory performance. As above, if the first block encountered by the participant is Pre-Exposure, it is suggested that overcoming the experimenter-introduced familiarity in this condition led to improvement in performance in the second block: No Pre-Exposure. This improvement in performance across these blocks is reflected in the greater use of Remember responses in the No Pre-Exposure condition. Usually the pattern of use of Remember responses would be the reverse of this with more Remember responses made in Pre-Exposure, as was the finding in Experiment 3.1. This was also the pattern found if the conditions were performed in the opposing order. If Pre-Exposure followed No Pre-Exposure the patterns were reversed and a greater proportion of items were assigned to Remember, and less to Familiar, in the Pre-Exposure condition. In this order, although no differences were found in memory performance, participants were changing their use of subjective experience categories in the expected direction to reflect the changing demands of the task: they were using recollection to a greater extent to overcome the exposure-induced familiarity.

This effect of block order is similar to that observed by Bodner and Richardson-Champion (2007) who found that when detail difficulty was manipulated across blocks the level of difficulty of one block influenced the subjective experience responses given to items in the next block. However, Bodner and Richardson-Champion also found that informing participants of the relative difficulty of the upcoming block eliminated this effect. Further experiments using the current pre-exposure paradigm could explore whether informing participants about whether a set of items are going to be pre-exposed, or whether that pre-exposure is going to be of irrelevant (distractor) words or target and lure words, changes their patterns of subjective experience responses. If informing participants about the nature of the pre-exposure was found to reduce the effects of blocking this would imply that the strategic use of recollection to overcome familiarity is able to be consciously controlled. However, the overall findings from these two experiments suggest that the differences in relative familiarity between pre-exposed and non-pre-exposed items were too small to be perceived by participants, thus it seems unlikely that people would be able to change how they interpreted this familiarity in light of experimental instructions.

One difference found between the results of the current experiment and Experiments 3.1 and 3.2 was that here participants assigned the same proportions of responses to Remember and Know. Although the means for proportion of responses assigned to Remember in Figure 3.6 appear to be much higher than those for Know, the large standard errors mean that these differences did not reach significance. The use of Familiarity in both tasks was also reduced in this experiment. It is suggested that perhaps the different methodology in Experiment 3.3 – use of Yes/No and non-associative 2AFC recognition led to these changes and future studies could explore this further by employing Yes/No and non-associative 2AFC recognition in between-subjects manipulations of pre-exposure.

As discussed in Section 3.2.1, associative recognition is theorised to rely more on recollection than familiarity as both target and lure items are familiar (in a re-arranged pairs design; Yonelinas, 1997, 2002; and Hockley & Consoli, 1999). However, in the three experiments reported here, higher levels of recollection (nearly .50) were found in Experiment 3.3 across both the item and non-associative AFC recognition tasks than in the associative-2AFC task used in Experiments 3.1 (Pre-Exposure condition) and 3.2, where only .39 and .36 of correct responses were assigned to Remember. Although this would appear to go against Yonelinas (1997, 2002) and Hockley and Consoli (1999), this finding is actually intuitive, here introduction of pre-exposure made the item and AFC tasks similar to an associative task: participants needed to rely on recollection as target and lure items were all familiar.

In comparing test formats, previous research had found mixed results with regard to levels of Remember responses in AFC and item recognition (e.g., Bastin & Van der Linden, 2003; Cook et al., 2005). Although not subject to statistical analysis, the results of the current experiment demonstrated no differences in use of subjective experience categories across test formats (see Figure 3.6) which fits with the results of Khoe et al., (2000) and Kroll et al., (2002).

A posteriori calculations of how likely an item was to have been recognised correctly if it was assigned to Remember, Know, Familiar, or Guess demonstrated similar, though not identical, patterns to those in Experiments 3.1 and 3.2, and patterns differed across tasks. In the 2AFC task, items assigned to Familiar showed the only significant difference, being more likely to be incorrect if in the Pre-Exposure condition. This latter result replicates results from Experiments 3.1 and 3.2 – when items were pre-exposed, using familiarity as a basis for recognition impaired performance. Conversely, in the Yes/No task, if an item was assigned to Remember or Know it was more likely to have been incorrect in the Pre-Exposure condition than in the No Pre-Exposure condition but the difference for Familiar responses was not significant. These results may be related to the increased use of Remember and Know responses in this task compared to the previous experiments. For Remember responses, as discussed above, the introduction of pre-exposure may have led participants to rely more heavily on recollection than is typical in item recognition tasks. Here this recollection has, in a significant proportion of cases, led participants to endorse lure items that participants only encountered in the pre-exposure phase as recognised from the study phase. If Know responses are considered to be high confidence responses experienced without recollection, as was found in Experiment 2.1 (Chapter 2), high levels of confidence from seeing an item in the pre-exposure phase may also have led to incorrect recognition. These false alarms to pre-exposed lures based on high confidence or recollection are more understandable in item recognition when there is the added confusion of pre-exposure as each item is assessed in isolation. Conversely, in 2AFC tasks recognition of one item is compared to that of another item and recollection of the target from the study phase should surmount any familiarity from seeing the lure in the pre-exposure phase.

3.4.4.3. Correlational analyses

Across both tasks and both exposure conditions memory performance was found to be positively correlated with proportion of correct items assigned to Remember and negatively correlated to proportion of items assigned to Familiar and Guess which replicates the correlations observed in Experiment 3.2. These data suggest that for all items in this blocked within-subjects design, pre-exposed or not, participants used recollection to resolve the familiarity induced by pre-exposure and this led to improved performance.

Interesting differences were observed in relation to the strength of the correlations in the two conditions. In the Yes/No task the positive correlation between memory performance and proportion of responses assigned to Remember was larger in the No Pre-Exposure condition than in the Pre-Exposure condition. This difference appears counterintuitive as use of recollection to overcome familiarity induced by pre-exposure should aid performance more in the Pre-Exposure condition. However, this result can easily be understood in terms of carry-over effects. Participants assigned the same proportion of responses to Remember in both conditions but as they did not 'need' to use recollection in the No Pre-Exposure condition, as there was no familiarity to overcome, this led to improved memory performance and a stronger relationship between memory performance and use of recollection. Conversely, in the 2AFC task the negative correlation between memory performance and proportion of responses assigned to Familiar was larger in the No Pre-Exposure condition than in the Pre-Exposure condition. This difference is not so easy to explain theoretically. As proportion of responses was so low in both conditions – around .15, and both correlation coefficients are over .60 it is suggested that the significant difference is an anomalous result. In reality, using familiarity as a basis for recognition, although not done to a great extent, impaired recognition in both conditions.

3.4.4.4. Reaction time analysis

In general the RT results demonstrated here matched those obtained in Experiment 3.1. Recognition decisions which were classified as Remember were made more quickly than other recognition decisions replicating Dewhurst et al. (2006) and Henson et al. (1999). However here there were differences between the tasks. In Yes/No, Remember and Know decisions were made significantly more quickly than Familiar decisions, a replication of patterns from Experiment 3.1; whereas in 2AFC Remember responses were faster than Know responses, which in turn were faster than Familiar responses. This different pattern could be because more Know responses, and fewer Familiar responses, were made in this experiment compared to in Experiment 3.1. For a full discussion of these findings see Section 3.5.3 in the General Discussion of this chapter.

3.5. GENERAL DISCUSSION

3.5.1. Memory performance

In all three experiments, the manipulation of pre-exposing participants to some/all of the stimuli that would be encountered during the study and test phases of the experiment led to impaired memory performance for items that had been pre-exposed. This replicates previous findings by Dobbins, Kroll, Yonelinas, et al. (1998), Kormi-Nouri et al. (2005), Åberg & Nilsson (2001, 2003), Tulving and Kroll (1995), Maddox and Estes (1997), Chalmers and Humphreys (1998), and Greene (1999). More importantly, the experiments reported here demonstrated that this memory performance effect is consistent across between-, within-, and blocked within-subjects designs and across item and 2AFC recognition.

3.5.2. Subjective experience

The *a priori* analysis of subjective experience responses demonstrated that the proportion of correct responses that were assigned to each category differed depending on the design of the experiment. Only in Experiment 3.1, which used a between-subjects methodology, were differences in the patterns of use of subjective experience categories observed. More responses were assigned to Remember and fewer to Know when items had been pre-exposed. Taken together with the memory performance results above, results from Experiment 3.1 demonstrate a case, when items have been pre-exposed, where memory performance is reduced but proportion of Remember responses increases.

In Experiments 3.2 and 3.3, high levels of Remember responses were observed across both pre-exposed and non pre-exposed items which is suggested to be due to carry-over effects between conditions (Conway et al., 2001). Here participants could not tell which items had been pre-exposed and used recollection to determine where they had encountered both pre-exposed and not pre-exposed items. Carry-over effects and the importance of experimental methodology in researching subjective experience are returned to later (Section 3.5.4). Comparing patterns of responses across the three experiments, other slight differences were observed, such as the same proportion of responses being assigned to Remember and Know and fewer responses being assigned to Familiar in Experiment 3.3. It is suggested that the different types of recognition tasks employed in Experiment 3.3 led to these changes. Future research could explore this by combining the between-subjects manipulation of Experiment 3.1 with the different types of recognition test employed in Experiment 3.3.

In Experiment 3.1 where differences in subjective experience were found, the patterns of results do not fit with earlier findings and conclusions drawn by Tulving and Kroll (1995) and Åberg and Nilsson (2003). Both these earlier studies asked participants to make confidence judgments on a 3-2-1 scale, hypothesised to map onto Remember-Know-Guess judgments (Åberg & Nilsson, 2003). Their results showed higher confidence for correct decisions when items *had not* been pre-exposed. If a high confidence response is thought to equal a Remember response then this pattern lies in direct opposition to the pattern observed in Experiment 3.1 where more Remember responses were obtained for items that *had* been pre-exposed. As was discussed in Chapter 2 (and is one of the focuses of Chapter 4), although confidence and Remember-Know judgments are no doubt related, this example highlights how the two can produce very different patterns of responses.

Åberg and Nilsson (2003) discussed their results in terms of recollection as they performed separate analysis just on high-confidence responses. Here they found increased correct recognition for non-pre-exposed (novel) items and suggested that the novelty effect (Tulving et al., 1994, 1996; Tulving & Kroll, 1995) is reliant on recollection: increases in performance for novel items are only evident when recollection is the basis for recognition. Results of the *a posteriori* analysis of Experiment 3.2 support their conclusion. In this analysis, likelihood of a response being correct was higher in the No Pre-Exposure condition than in the Pre-Exposure condition if the response was assigned to Remember. With the within-subjects manipulation of pre-exposure in Experiment 3.2, recollection of items from encoding benefitted memory performance for novel (non-pre-exposed) items to a greater degree than it did for pre-exposed items. However, the *a posteriori* analysis for Experiment 3.3 demonstrated a different pattern of findings. In the Yes/No task the likelihood of a recognition decision being correct was higher in the No Pre-Exposure condition than in the Pre-Exposure condition for responses that were assigned to either Remember or to Know. Here it was not only recollection which aided recognition, knowing also improved performance. In contrast, in the 2AFC task of Experiment 3.3 there were no differences between conditions for the likelihood of Remember and Know responses being based on correct recognition decisions. Thus, these findings do not show a consistent pattern of support for Åberg and Nilsson's suggestion that the novelty effect is reliant on recollection.

In general, the results of the current experiments are somewhat difficult to interpret in terms of the novelty-encoding hypothesis (Tulving et al., 1994, 1996; Tulving & Kroll, 1995). One of the issues which has been debated between researchers is whether the novelty effect is evidenced in correct recognition or in false alarms (Dobbins, Kroll, Yonelinas, et al., 1998; Åberg & Nilsson, 2003). If differences are only found in false alarms then it is not novelty that aided recognition, it is pre-exposure that impaired it (Åberg & Nilsson, 2003).

However, as discussed previously, false alarms are different in item and 2AFC recognition. In 2AFC recognition a false alarm is not simply an endorsement of a lure item, it is endorsement of a lure item over a target item. In all experiments presented in this chapter, memory performance was more accurate when items had not undergone pre-exposure. The question is whether this was due to the novel, non-pre-exposed, items being encoded better – as predicted by the novelty-encoding hypothesis, or false alarms occurring for pre-exposed items – due to problems of source discriminability (Dobbins, Kroll, Yonelinas, et al., 1998).

Experiment 3.1 cannot directly test the novelty-encoding hypothesis as using the between-subjects design all the items a participant studied were either novel (in the No Pre-Exposure condition) or were all familiar (in the Pre-Exposure condition). Within condition, there were no differences of novelty/familiarity at encoding. However, there were differences in novelty/familiarity at retrieval in both conditions. In the No Pre-Exposure condition, the target items had been studied and were therefore more familiar than the lure items, which were completely novel. In contrast, in the Pre-Exposure condition, the lure items had been pre-exposed and therefore had some level of familiarity, however, the target items had been pre-exposed *and* studied, therefore they should have been more familiar to participants than the lures. Examination of the *a posteriori* data in Experiment 3.1 shows how these relative differences in familiarity affected subjective experience judgments and memory performance. The only difference that was observed was that the likelihood of recognition being correct was lower in the Pre-Exposure condition than in the No Pre-Exposure condition when responses were assigned to Familiar. This can be interpreted as pre-exposure leading to impaired recognition due to more false alarms being based on familiarity and thus supports the source-discriminability hypothesis. When basing a recognition decision on familiarity participants found it hard to determine where an item was familiar from – *had it been studied or had it only been encountered on the pre-exposure task?*

In contrast to Experiment 3.1, Experiment 3.2 did include differences in the relative novelty/familiarity of items at encoding. When pre-exposure was manipulated within-subjects some items were already familiar when encountered in the study phase (as they had been pre-exposed) whereas some items were novel when presented for study. Using this manipulation, the *a posteriori* analysis revealed that both Remember and Familiar responses were less likely to have been made to a correct recognition decision for pre-exposed compared to non-pre-exposed items. These findings can be interpreted in two ways. Pre-exposure could have led to impaired recognition due to more false alarms being based on familiarity or recollection of the item from the pre-exposure phase; or, the novelty of non-pre-exposed items when first encountered in the study phase could have led to deeper processing at encoding and enhanced recollection and familiarity at retrieval. Thus, both source-discriminability and novelty-encoding explanations can account for the pattern of findings observed.

Across experiments, analysis of correlations between memory performance and subjective experience responses demonstrated similar patterns of findings as the *a priori* analysis. For Experiment 3.1, the *a priori* analysis demonstrated higher levels of Remember responses in the Pre-Exposure condition compared to the No Pre-Exposure condition. The correlational analysis revealed a positive correlation between Remember responses and memory performance and a negative correlation between Familiar responses and memory performance but only for the Pre-Exposure condition, in the No Pre-Exposure condition no type of subjective experience correlated with memory performance. However, in Experiments 3.2 and 3.3 these differences between conditions were not observed. Patterns of use of subjective experience responses were the same across conditions and memory performance was positively correlated with use of Remember responses and negatively correlated with use of Familiar responses whether items had been pre-exposed or not. As discussed previously, these results can be explained by carry-over occurring from one condition to another. In Experiments 3.2 and 3.3 participants could not tell which items had been pre-exposed and used recollection to determine where they had encountered both pre-exposed and not pre-exposed items.

Taken together, results from all three experiments suggest that distinguishing between novelty-encoding and source-discriminability accounts of how pre-exposure impairs memory cannot be formally tested by examining subjective experience. Testing the novelty-encoding hypothesis requires a within-subjects manipulation of pre-exposure so that there are differences between novelty of items when they are being encoded. However, manipulating pre-exposure within-subjects was found to result in carry-over of use of subjective experience from one condition to the next. Participants were not able to determine which items have been pre-exposed and therefore which they would *need* to use recollection to recognise so they used recollection as the basis for all recognition decisions. However, the very fact that this carry-over occurs suggests that the memory impairment found with pre-exposure is down to the retrieval of items from memory, not from their encoding, therefore supporting the source-discriminability hypothesis of Dobbins, Kroll, Yonelinas, et al. (1998).

This suggestion also fits with the SAC model proposed by Reder et al. (2000; Cary & Reder, 2003) to explain word frequency effects. In their model of spreading activation between word nodes and event nodes, increases in false alarms for high-frequency words result from higher base activation of their word nodes. False alarms occur because participants accept high-frequency words as old because they have misattributed their pre-experimental familiarity. For the current data, the emphasis in this model would simply have to be changed from word nodes to event nodes – from pre-experimental familiarity to experimentally manipulated familiarity. In the experiments presented here differences in memory performance occurred because participants misattributed experimental familiarity from the pre-exposure phase to the study phase.

3.5.3. Reaction time

Reaction times were found to be faster for Remember responses in both Experiment 3.1 and 3.3 replicating previous findings (Dewhurst & Conway, 1994; Dewhurst et al., 1998, 2006; Henson et al., 1999; Vilberg & Rugg, 2007; Wixted & Stretch, 2004), though in the current experiments in general the difference was found between Remember and Familiar responses, not Remember and Know. These findings go against assumptions that recollection is a slower and more effortful process than familiarity (e.g., Jacoby & Dallas, 1981; Mandler, 1980).

The current results demonstrate that recognition decisions based on recollection, and measured by Remember responses, can be made quickly and accurately. In contrast, RTs found here for Know and Familiar responses demonstrate that these may require additional post-retrieval processes to establish their familiarity relative to other items on the test (Dewhurst et al., 2006), or further attempts at retrieval (Henson et al., 1999), may be required. This suggestion is particularly poignant here as RTs to Know and Familiar responses showed different patterns between the two experiments. Know responses were as fast as Remember responses in Experiment 3.1 and the Yes/No task of Experiment 3.3, and both were faster than Familiar responses; however in the 2AFC task of Experiment 3.3 Remember were faster than Know which in turn were faster than Familiar. These reaction time findings could be associated with the differences in confidence associated with Know and Familiar responses (see Chapters 2 and 4). In the current experiments, high levels of confidence could have led to equally fast responding for Know responses as for Remember responses, whereas the lower confidence associated with Familiar responses may have led to them requiring additional processing and being made more slowly. This idea was explored experimentally by Rotello and Zeng (2008). Their results are discussed in detail in the Introduction of Chapter 4 (Section 4.1.3) as the relationship between judgments of subjective experience and varying levels of confidence is explored through reaction time analysis in the experiments in that chapter.

3.5.4. Experimental design for subjective experience experiments

One of the major conclusions to be drawn from the three experiments here is that in some cases differences in subjective experience can only be explored via between-subjects comparisons. Previous research has found the opposite to also be true in some cases; for example, Dewhurst and Parry (2000) found that effects of emotionality on Remember and Know judgments was only demonstrated using within-subjects manipulations. However, the findings of Conway and Dewhurst (1995a), the ‘non-replication’ by Hirshman and Lanning (1999) and re-replication by Conway et al. (2001) demonstrate how small differences in experimental manipulations can influence results when it is subjective experience you are interested in. It is always beneficial to explore relationships using multiple experimental designs. For the results of the experiments presented here this is particularly important as using the within-subjects designs the memory performance impairment for pre-exposed items was still demonstrated, but differences in patterns of subjective experience were not.

Using within-subjects designs here demonstrated evidence of carry-over of use of subjective experience categories from the Pre-Exposure condition to the No Pre-Exposure condition. This carry-over was evidenced in the *a priori, a posteriori,* and correlational analyses. Examination of effects of block order in Experiment 3.3 demonstrated this contamination clearly, as use of subjective experience categories in the second block was dependent on which condition had been experienced in the first block, and these changes in use of subjective experience were found to influence memory performance in the two blocks also. The differences in patterns of subjective experience responses observed across between- and within-subjects manipulations of pre-exposure are akin to the findings of Whittlesea (2002). Comparing across all three experiments in this chapter, the different contexts in which pre-exposed and non-pre-exposed items were encountered influenced how those items were subjectively experienced. As Whittlesea (2002) stated, “the specific phenomenology that a person experiences depends on his or her interpretation, within some specific context, of the significance of those aspects of performance which are salient within the event” (p. 325). As discussed previously (Section 3.4.4.2), in future research it would be interesting to examine whether explicit instructions about blocking (Bodner & Richardson-Champion, 2007) or about the relationship between pre-exposed and study items would remove the carry-over effects between blocks in a blocked design, or between items in a within-subjects design.

3.5.5. Conclusions

On a methodological note these experiments demonstrate the importance of using between-subjects designs to explore subjective experience. On a theoretical level they demonstrate that when familiarity is manipulated in such a way a situation occurs where increases in recollection are associated with reductions in memory performance which has implications for cognitive theories of recollection. The use of recollection to overcome the familiarity induced by pre-exposure also demonstrates the independence of the processes of recollection and familiarity (Dobbins, Kroll, Yonelinas, et al., 1998, cf., Jacoby, 1991).

In regard to a key theme of this thesis, the three experiments in this chapter provide further evidence that Know and Familiar can be separated as categories of subjective experience. The proportion of responses assigned to Know was found to be affected differently by pre-exposure in between- and within-subjects manipulations of pre-exposure while the proportion of responses assigned to Familiar responses was unaffected by pre-exposure and remained the same across the different experimental designs. In contrast, Familiar responses were shown to differ in terms of accuracy across pre-exposure conditions and were shown to have a relationship with memory performance when items had been pre-exposed. Finally, Know and Familiar responses were also found to differ in terms of reaction time and accuracy; Know responses being faster and more accurate than Familiar responses. The relationships between accuracy, reaction time, and subjective experience are further explored in Chapter 4.

Results of the current experiments are also relevant to the question of how confidence and subjective experience are related. As discussed previously, results of Experiment 3.1 demonstrated an opposite pattern of findings to those observed by Tulving and Kroll (1995) and Åberg and Nilsson (2003). These earlier studies found higher confidence for correct decisions when items *had not* been pre-exposed. In contrast, in Experiment 3.1 more Remember responses were obtained for items that *had* been pre-exposed and in Experiments 3.2 and 3.3 there were no differences in proportion of items assigned to Remember across conditions. Thus, comparing these findings demonstrates that confidence and subjective experience do not respond in the same manner to manipulations of pre-exposure. The relationship between confidence and subjective experience is a central focus of the experiments presented in Chapter 4.

4. Subjective Experience, Confidence, and Source: Differences, Influence, and Reaction Time

"As remarked by an anonymous reviewer, whether confidence ratings can safely accompany an R-K(-G) judgement, and whether they should follow or precede such judgement, is as much an empirical as it is a theoretical question."
(Bruno & Rutherford, 2010, footnote 2, p. 127).

In this chapter three experiments are presented which examine the influence that judgments of confidence, source, and subjective experience have on each other. Analysis focuses on distribution of responses across confidence levels, subjective experience categories, and source accuracy, dependent on whether judgments were made first or second following a recognition decision. Reaction times to make Old/New recognition responses are analysed dependent on subsequent judgment and, in a novel approach to this field, the time taken to make the actual judgments is also explored. The experiments presented here lead on from those in Chapter 2 which showed reliable differences between the four categories of subjective experience (Remember, Know, Familiar, and Guess) with regards to confidence. In the present chapter, these four categories of subjective experience are contrasted against four levels of confidence. These experiments also lead on from those in Chapter 3 which demonstrated that Know responses were as accurate and made as quickly as Remember responses.

4.1. General introduction

As discussed in Chapter 1, the relationship between confidence and subjective experience is of critical importance to the debate between single- versus dual-process accounts of memory (e.g., Dunn, 2004, 2008). If the two types of judgment produce the same patterns on a task then this suggests that the two are tapping the same underlying memory processes and thus would support a single-process account of memory. In contrast, if different patterns of findings are observed then this suggests that there is more than the one process of trace strength underlying memory decisions and would thereby provide support to dual-process accounts of memory (e.g., Gardiner, 2008). The current experiments aim to gain a better understanding of subjective experience theoretically by comparing the subjective measures of Remember-Know and confidence with the objective measure of source accuracy.

As stated in the quote at the start of this chapter, the issue of whether judgments influence each other when made together is of interest both empirically and theoretically. However, many experimental paradigms combine judgments without the issue of influence having been explored. Experiments that have compared combinations of two of these three judgments are discussed in the following sections. In addition to focusing on response patterns for subjective experience, source, and confidence judgments when different judgments are compared, experiments which have explored reaction times associated with different responses are also reviewed. Reaction time to make judgments was a primary analysis of interest in the current experiments as, to my knowledge, no one has yet compared how long it takes participants to make these three different types of judgment.

Within each of the following sections, methodological differences pertinent to the design of the current experiments are discussed. For example, one methodological issue in the Remember-Know paradigm is whether use of one-step or two-step Remember-Know procedures influences recognition accuracy or makes Remember responding more lenient. Bruno and Rutherford (2010), Eldridge et al. (2002), and Hicks and Marsh (1999) have explored this issue in depth and their findings were discussed in the General Introduction (Section 1.6.5.1). However, exploration of source, confidence, and subjective experience judgments using one-step and two-step procedures has produced some differing findings and methodological differences are therefore highlighted. In addition, the experimental design of the studies presented here, with two judgments being made post-recognition, is of particular importance for patterns of influence and reaction time analysis.

4.1.1. Comparing subjective experience and confidence judgments

As discussed in Chapter 2, Tulving (1985) originally demonstrated that Remember judgments are typically associated with higher confidence ratings than Know judgments. This finding has been widely replicated (e.g., Brewer & Sampaio, 2006) and in Chapter 2 Experiment 2.1 yielded the same finding with the novel task of rating others' recognition memory justifications. These studies were interested in confidence ratings for items categorised as Remember or Know but other studies have explored the equivalence of subjective experience judgments and confidence judgments by comparing them in experimental paradigms. As discussed in depth in the General Introduction (Section 1.6.4), traditionally confidence was operationalised as a two-category scale of Sure-Unsure and proportion of items assigned to these two categories were compared against proportion assigned to Remember and Know. Gardiner and Java (1990) thus compared subjective experience and confidence to words and non-words and demonstrated that these two measures did not respond equivalently to word and non-word recognition. More Remember judgments were given to words while more Know judgments were given to non-words. This interaction between word type and response was not found when Sure-Unsure judgments were made. These patterns were replicated in similar studies by Rajaram (1993), Rajaram et al. (2002), and Tunney and Fernie (2007) and together they demonstrate that while judgments of confidence and subjective experience may be interrelated, people do not use the judgment types in the same way.

Analysis of subjective experience across differing levels of confidence has also been examined using ROC curves with the consistent finding that higher confidence is associated with Remember responses compared to Know responses (Rotello et al., 2004, 2005; Slotnick, 2010; Wixted & Stretch, 2004; Yonelinas, 2001a; Yonelinas et al., 1996). However, as discussed in Chapter 2 (Section 2.1.1), distribution of Remember and Know responses across confidence levels has been found to be influenced by how conservative the Remember-Know instructions are (Rotello et al., 2005; Slotnick, 2010). Whilst these results are interesting for the single- versus dual-process debate, the manner in which confidence and subjective experience are operationalised in ROC studies means that participants' response distributions for the two cannot be directly compared as was done when Sure-Unsure were compared against Remember-Know.

To foreshadow Experiments 4.2 and 4.3, source monitoring is another method used to explore recollection processes (see the General Introduction, Section 1.5.1.2). Experiments that have compared source judgments with confidence and/or subjective experience judgments are discussed after Experiment 4.1.

4.1.2. The influence of one judgment on another judgment

To date, no studies have explored the influence that confidence and subjective experience judgments may have on each other when the two judgments are reported post recognition. The traditional viewpoint on this subject is summarised by Holmes, Waters, and Rajaram (1998) who stated that "…multiple judgments might affect performance in unforeseen ways" (p. 1031) and Humphreys et al. (2003) who thought that "When two judgments are made on the same trial, the knowledge that a second judgment will be required could distort the first judgment" (p. 806). This issue has also recently been considered by Bruno and Rutherford (2010). These authors chose to not combine confidence and subjective experience judgments in their experiment:

> It is our belief that collecting confidence ratings when also R–K(–G) judgements are provided is problematic because these measures may not be independent of each other. R–K(–G) judgements may directly influence confidence ratings (e.g., 'if I *remember* an event, then I must assign a higher rating') and, vice versa, confidence ratings may exert an influence on the R–K(–G) judgement (e.g., 'if I give a high rating, then it must be *remember*'). (italics in original; Bruno & Rutherford, 2010, p. 127)

These comments highlight that this belief has existed in the Remember-Know paradigm across at least the last two decades but has not yet been subject to empirical testing.

In line with the traditional viewpoint of Holmes et al. (1998), previous experiments comparing subjective experience and confidence judgments have employed measures to ensure that the two judgment types did not influence one another. In the original subjective experience and confidence judgment comparisons by Gardiner and Java (1990) and Rajaram (1993), judgment type was compared across two separate experiments. Rajaram et al. (2002) later replicated these findings using a within-subjects design, but even in this design testing sessions employing the two types of judgments took place one week apart and subjective experience was always tested first, "…to ensure that confidence judgments did not contaminate the remember-know judgments with a carryover effect to the second session" (p. 230). So although the same participants were making both types of judgment, Rajaram et al. were still worried about how one judgment might influence the other.

4.1.3. Reaction time data associated with confidence and subjective experience judgments

As with the influence of judgments on each other, reaction times to make post-recognition judgments have not before been tested experimentally. In two-step procedures, when examining reaction times to items recognised with differing levels of confidence or subjective experience the standard analysis is to compare recognition decision RTs for items categorised by their subsequent judgment category. In one-step procedures it is the judgment of subjective experience or confidence which is timed, but this judgment is made in conjunction with the recognition decision so the timings of the two decision processes are indistinguishable. However, reliable differences have been demonstrated for both judgment types using both these types of RT analysis.

Recognition is faster when confidence is higher[19]. Mandler and Boeck (1974) used a two-step Yes/No then 3-point confidence scale procedure and found recognition decisions to highest confidence items to be approximately 1200ms faster than recognition decisions to items given the lowest confidence ratings. Using a one-step 6-point confidence recognition scale this pattern was replicated by Ratcliff and Murdock (1976) who demonstrated a 1000ms (approx.) difference between highest and lowest confidence RTs. Recent evidence from neuroimaging has also produced similar findings. For example, Henson et al. (2000) demonstrated an advantage in RT of approximately 600ms for high confidence compared to low confidence recognition judgments.

Recognition RTs are also faster for Remember items than for Know items (Dewhurst & Conway, 1994; Dewhurst et al., 1998, 2006; Henson et al., 1999; Vilberg & Rugg, 2007; Wixted & Stretch, 2004). Dewhurst et al. (2006) demonstrated this finding to be true when both one-step and two-step procedures are used, and also found reliable differences in recognition RTs when stimulus characteristics are varied; for example a larger RT advantage for Remember responses for words compared to non-words. However, differences in Remember and Know recognition RTs can be much smaller than those observed for confidence, for example Wixted and Stretch (2004) reported a difference of only 80ms to be reliable. In comparison, other researchers have found as large an RT

[19] While RTs for incorrect recognition decisions are worthy of note (cf. Wixted & Stretch, 2004), for the purposes of this chapter discussion will focus on RTs for correctly recognised items.

advantage for Remember over Know responses as that obtained for high compared to low confidence; for example a 730ms advantage was reported by Dewhurst et al. (2006) and an 830ms advantage by Henson et al. (1999).

The experiments presented in Chapter 3 of this thesis also demonstrated that recognition RTs were faster for correctly recognised items then assigned to Remember, though with the four categories of subjective experience that were employed in these experiments (Remember, Know, Familiar, and Guess), in general the difference in RT was found to be between Remember and Familiar responses, with no differences in RTs for Remember and Know (except Experiment 3.3 which did find a difference between Remember and Know RTs for the 2AFC task). These results fit with the patterns observed by Rotello and Zeng (2008). They found Remember to be faster than Know responses, however participants in their experiments also made confidence judgments and when Remember and Know responses were equated for confidence the RT advantage for Remember responses was reduced. However, in the experiments presented in Chapter 3, whereas confidence judgments were not made, Remember and Know responses were equated for accuracy. No differences in recognition RTs or accuracy for Remember and Know responses was demonstrated, but Familiar responses, which had lower levels of accuracy, also had longer recognition RTs. This issue will be returned to in the General Discussion of this chapter (Section 4.4.3).

So why have judgment RTs not been explored before? One methodological reason is that the participant often has to wait till a prompt appears on the screen before they can make their judgment (Wilding & Rugg, 1996). The assumption perhaps being that the participant would have decided on their response prior to the prompt and therefore no difference in RTs would be observed. In the experiments presented in this chapter this problem was avoided through counterbalancing of judgment order and the inclusion of 'catch' trials where participants were asked to make other non-critical judgments for the item. On these trials, the first judgment made by participants was one critical to the experiment (e.g., either confidence or subjective experience in Experiment 4.1); however, the second judgment was either a judgment of age of acquisition (AoA) or of pleasantness. As participants would not have known which judgment they would be required to make first post-recognition, and they certainly would not know what type of judgment would be coming second, the issue of pre-prompt decisions being made for judgments is reduced. Participants could not systematically predict which judgment they would be making.

As the experiments presented here were interested in comparing two judgments made post-recognition, a methodological reason for analysing judgment RTs in addition to recognition RTs classified by subsequent judgment is that both judgments made for a single item are dependent on the same recognition decision and therefore this RT could not be used to compare judgment types across judgment order. A theoretical reason for analysing judgment RTs is that whilst recollection and high confidence have each been shown to increase the speed at which a separate recognition decision is made, research has not yet examined whether these can influence RT for decisions which are independent of recognition. This separation of recognition decision time and judgment decision time also enables comparison of judgment types when employed in identical paradigms as in the literature RTs vary between studies. Here RT to make subjective experience and confidence judgments can be directly compared for the first time. It is predicted that confidence judgments will be made more quickly than judgments of experience as it is suggested that the consideration of whether anything is recollected about the encoding of the item will take longer than simply thinking about the confidence associated with recognition. RT for judgments is also interesting from an influence point of view. In addition to examining whether making a Remember judgment always leads to a secondary judgment of high confidence, for example (Bruno & Rutherford, 2010), RTs for judgments can explore whether making a Remember judgment also influences the speed at which that later confidence judgment is made, and vice versa.

The current experiment was designed to explore whether when two judgments were made post-recognition the response to one judgment influenced the other judgment. The examination of influence of subjective experience and confidence judgments on each other is entirely novel, as is examination of RTs to make judgments that occur after, and are independent of, recognition. The three experiments presented here are identical in procedure except that the two post-recognition judgments differ across experiments. In Experiment 4.1 confidence judgments are compared with judgments of subjective experience.

4.2. EXPERIMENT 4.1: SUBJECTIVE EXPERIENCE AND CONFIDENCE JUDGMENTS

4.2.1. Introduction

In this experiment participants studied medium-frequency words and at test made an Old-New recognition judgment followed by two post-recognition judgments. The two critical judgments made in this experiment were confidence and subjective experience. For both judgment types four response options were provided: Remember, Know, Familiar, and Guess for subjective experience judgments, and 3 (High), 2 (Medium), 1 (Low), and 0 (None) for confidence judgments. The order in which participants made these judgments was counterbalanced across items.

On the basis of previous findings it was predicted that Remember judgments would be associated with higher levels of confidence than Know, Familiar, and Guess judgments (e.g., Yonelinas, 2001a; Yonelinas et al., 1996) but that Know judgments would have as high recognition accuracy and as fast a recognition reaction time as Remember judgments (Rotello & Zeng, 2008; Experiments 3.1 and 3.3 of this thesis). It is assumed that the two judgment types are not “experimentally interchangeable” (Rajaram et al., 2002, p. 234) and to examine this, distribution patterns, accuracy, and RT data for the four categories of subjective experience and confidence are explored. If confidence and subjective experience are independent then it would be expected that a fully counterbalanced design would reveal dissociations between the two types of judgment on these key dependent variables.

With regard to the influence of one judgment on the other, the order of judgments was explored. It is possible that Bruno and Rutherford’s (2010) suggestion is correct and that after making a Remember judgment participants will be more inclined to make a judgment of High confidence, for example; but comparing the order of judgments allows the reverse pattern to also be explored – whether more Remember judgments are made after a judgment of High confidence. Influence is also able to be explored in the RT analysis. As it is considered that subjective experience gives rise to confidence, not the other way around (Gardiner, 2001), it is suggested that confidence judgments that are made following a judgment of subjective experience may be speeded as assessment of recollection and familiarity has already occurred when the subjective experience judgment was made.

4.2.2. Method

4.2.2.1. Participants and Design

Participants were 44 undergraduate students (38 female), mean age 19.25 years (range 18 to 23), from the University of Leeds, who received either Participant Pool Credits or £5 for taking part. Participants were tested in groups of between three and fourteen at individual PCs. Data from two further participants could not be analysed due to computer problems during testing. The experiment employed a within-subjects design and for the study and test phases all instructions and stimuli were presented, and data collected, using E-Prime version 1.2.

4.2.2.2. Materials

Information concerning materials and procedures are presented here in detail for all experiments in this chapter. Critical methodological differences are then discussed for Experiments 4.2 and 4.3. The target words were 56 medium-frequency words[20] (mean familiarity rating of 424; range 350-480) limited to between five and eight letters in length. Targets were matched with 56 words to be used as lure items in the recognition test, and 16 filler words. In total participants studied 64 words (56 targets plus 8 fillers, 4 shown at the start and 4 at the end of the study phase) and performed the recognition test on 128 words[21]. This 128 included the 8 studied fillers and a further 8 lure fillers; all fillers were shown at the start of the test phase to acquaint the participants with the procedures for making recognition decisions and judgments. All fillers were excluded from analysis.

Participants made four types of judgment in this experiment: subjective experience judgments (Remember, Know, Familiar, Guess), confidence judgments, age-of-acquisition (AoA) judgments, and pleasantness judgments. Participants were always required to make

---

[20] Words obtained from the MRC Psycholinguistic Database. Familiarity values refer to the printed frequency in the language and were derived from merging three sets of familiarity norms: Pavio (unpublished), Toglia and Battig (1978) and Gilhooly and Logie (1980).

[21] After data sorting it was discovered that the word ‘wiggle’ had been used as both a target and a lure item. All data for this item were therefore deleted and analysis of target and lure lists consisted of 55 items each. This was done for all three experiments presented in this chapter.

either a subjective experience judgment or a confidence judgment first, followed by another type of judgment (of the three not already made). The critical trials were those where participants made a combination of a subjective experience judgment and a confidence judgment, either subjective experience then confidence or confidence then subjective experience. The AoA and pleasantness trials were included in an attempt to avoid participants guessing the manipulation of interest and so that participants were not able to predict what judgment would follow the first. Assigned to the 64 studied words (56 targets, 8 fillers) were 16 trials of each judgment pairing, either: subjective experience + confidence, confidence + subjective experience, subjective experience + AoA/pleasantness (eight of each), and confidence + AoA/pleasantness (eight of each). Words were randomly assigned to judgment pairings, though it was ensured that filler words were always followed by a non-critical judgment pairing (involving AoA or pleasantness). Lure words were assigned to judgment pairings in an identical manner to target words. AoA and pleasantness responses were not analysed.

Four word lists were utilised. Two sets of 64 words were counterbalanced as lure and target lists across participants. Two versions of each were constructed where each word was matched with a different judgment pairing and which also had differing inter-trial-intervals (ITI) and inter-judgment-intervals (IJI) for each word in the test phase. Allocation of participants to list was random.

4.2.2.3. Procedure

In the study phase, target words were presented in random order individually in the centre of the computer screen. Participants were instructed that they should study the words and that later they would undergo an Old-New recognition test. It was emphasized that it was a long list of words and that participants should concentrate on the screen throughout. Participants studied each word for four seconds separated by a fixation point shown for 750 milliseconds. The eight filler words were shown in random order at the start and the end of the target word list. Prior to undergoing the recognition test, participants completed the pattern comparison and letter comparison speed of processing tasks (e.g., Salthouse, 1991). Although not formally timed, participants took approximately five minutes to complete this.

In the test phase, participants were instructed that for each word recognised as Old they would then be asked to make two judgments from the list of four judgments. The four judgments were then explained on individual screens (with full definitions of Remember (R), Know (K), Familiar (F), and Guess (G) provided on paper; see Appendix A.3). Each explanation screen included an example of the scale that the participant would use for each type of judgment. The subjective experience judgment scale consisted of four response boxes labelled (from left to right) R, K, F, and G, with the question 'What is your EXPERIENCE of recognising this word?'. The confidence scale consisted of four boxes labelled (from left to right) 0, 1, 2, 3, with 0 accompanied by the label 'Not confident at all' and 3 accompanied by 'Extremely confident', and the question 'How CONFIDENT are you that you correctly recognised this word?'. For ease of description in discussing analyses, these numerical confidence levels are translated into verbal labels: 3 = High, 2 = Medium, 1 = Low, 0 = None. The labelling of the subjective experience scale 'R, K, F, G' from left to right and the confidence scale '0, 1, 2, 3' from left to right was purposefully done to ensure that on trials involving both these judgments, if assigning the item to both Remember and High confidence for example, participants were not moving the mouse to exactly the same position to make both responses. In between judgments the mouse pointer was repositioned to the centre of the screen. The AoA judgment scale consisted of four boxes labelled 0-4, 5-8, 9-12, and 13+, and the question 'What AGE were you when you first learnt this word?'. Finally the pleasantness scale consisted of the question 'How PLEASANT is this word?' and four boxes labelled 1, 2, 3, and 4, with 1 accompanied by the label 'Not pleasant' and 4 accompanied by 'Very pleasant'. For each judgment type, the response boxes were the same size, shape and positioned identically.

Commencing with the 16 filler words, participants then underwent the recognition test. At the start of a trial a fixation point was shown for 750ms. Each word was then presented in the centre of the screen with the cues 'New' and 'Old' presented below and towards the left and right of the screen respectively, reflecting the number keys which participants were required to use to make their Old-New judgment: 1 for New, 2 for Old. Participants were instructed to press these keys with their left hand. Reminders for which numbers corresponded to Old and New were shown on all trials. When participants indicated that they recognised a word as Old the first judgment screen was shown. The word re-appeared in the centre of the screen with the response boxes and judgment question below. The mouse pointer appeared in the centre of the screen and participants made all judgments using the mouse to click on one of the response boxes. Speed and accuracy of responses were emphasised in the instructions for the test phase and response times for all recognition decisions and judgments were recorded. In between the two judgment screens and prior to a new trial commencing a blank screen was shown as an IJI or ITI for a duration which ranged between 250ms and 1250ms. Every twelve trials during the test phase participants were allowed to take a break if they were fatigued. A 'take a break' slide appeared and remained until the participant pressed the spacebar, on this screen was also a reminder that all responses should be made as quickly and as accurately as possible. After completion of the test phase participants completed a debrief questionnaire to ensure that they had understood the different types of judgments made; no participants were excluded on the basis of their responses.

4.2.3. Results

Memory performance was examined, followed by the proportion of correct responses assigned to each of the confidence levels and subjective experience categories calculated using both *a priori* and *a posteriori* methods. The influence of First Judgment on Second Judgment was then explored using *a priori* and *a posteriori* methods. Reaction times for recognition decisions and judgments were analysed by confidence level or subjective experience category. Prior to analysis, data from any responses which had been made faster than 300ms or slower than 8000ms were excluded from the dataset.

4.2.3.1. Memory measures

The mean proportion of targets correctly recognised (hits) was .62 (SD = .18) and the mean proportion of false alarms (FA) made to lures was .08 (SD = .08). Overall recognition performance was therefore .54 (hits minus FA; SD = .20) which was significantly higher than chance (zero), $t(43) = 17.86$, $p < .001$. Recognition was also examined using *d'*. The *d'* value obtained was 1.85 (SD = .77). As participants made FAs to only 8% of lures, analysis of lure data was not performed.

4.2.3.2. Subjective experience and confidence measures

Which levels of confidence or subjective experiences participants had based their recognition decisions on was the second analysis of interest. Firstly *a priori* proportions were calculated – the proportion of correct responses assigned to each of the confidence levels or subjective experience categories. These proportions sum to 1 for each participant; where this occurs it is noted and the variables are analysed together with planned comparisons used to further explore the data. As is shown in Figure 4.1, there were only slight differences between assignment of First and Second Judgments to different responses, but patterns of responses across confidence levels differed from patterns of responses across subjective experience categories. A 2(judgment type) x 2(judgment order) x 4(response category) within-subjects ANOVA was conducted on these proportions. ANOVA demonstrated a significant main effect of response category, $F(1.69,72.75) = 84.79$, $p < .001$, and a significant interaction between judgment type and response category, $F(1.43,61.61) = 13.61$, $p < .001$; participants did not use confidence levels in the same manner as subjective experience categories. Nearly 60% of responses were assigned to High confidence, while just over 40% of responses were assigned to Remember. Approximately equal proportions of responses were assigned to Know and Familiar, whilst slightly more responses were assigned to Medium compared to Low confidence. The proportions assigned to Guess and None were approximately equal and were low, at under 5%.

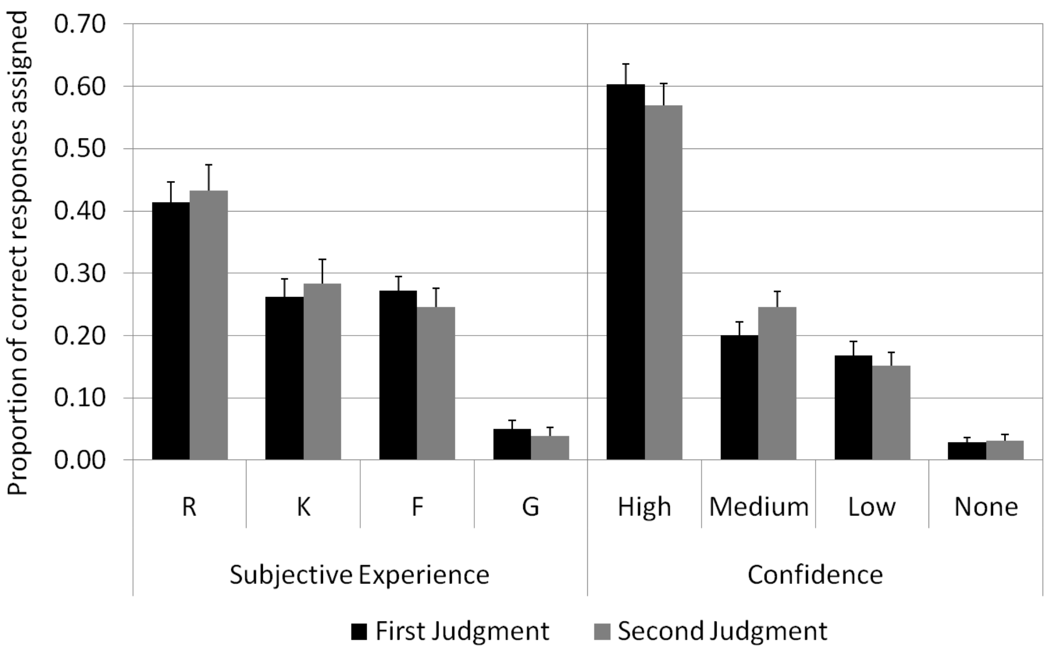

Figure 4.1. Mean proportion of correct responses assigned at First and Second Judgments to subjective experience categories Remember (R), Know (K), Familiar (F), and Guess (G), and confidence levels High, Medium, Low, and None. Errors bars = 1 SeM.

There was no interaction between judgment order and response category $F(1.97,84.86) = 1.49$, $p = .22$, the interaction between judgment type and judgment order was not calculated as proportions summed to 1, and the three-way interaction was not significant, $F < 1$. As there was no main effect of judgment order, judgment order did not interact with response category, and there was no three-way interaction, further analysis was conducted with data for First and Second Judgments combined together. Confidence and subjective experience do not appear to differ as a function of the other judgment being made before or after.

Aggregated across judgment order, separate 4(response category) ANOVAs for subjective experience and confidence were conducted to follow-up the significant interaction between judgment type and response category. The separate ANOVAs both demonstrated significant main effects of response category: subjective experience, $F(1.81,77.77) = 32.26$, $p < .001$, and confidence, $F(1.48,63.75) = 114.38$, $p < .001$. Planned comparisons were conducted between the different response categories for confidence judgments and judgments of subjective experience[22]. For subjective experience, planned comparisons across categories demonstrated there was no significant difference in the proportion of responses assigned to Know and Familiar, $t < 1$. All other comparisons were significant, all at least $p < .007$; more responses were assigned to Remember than to Know, Familiar or Guess, and fewer responses were assigned to Guess than to any other category. Comparisons across confidence levels demonstrated all proportions were significantly different, all at least $p < .006$. More responses were assigned to High confidence compared to Medium, Low, or None; more were assigned to Medium compared to Low confidence, and more were assigned to Medium or Low confidence compared to None. Planned comparisons across judgment type demonstrated that more responses were assigned to High confidence than to Remember, $t(43) = 4.87$, $p < .001$, and fewer responses were assigned to Low confidence than were assigned to Familiar, $t(43) = 6.35$, $p < .001$. Comparisons for Know against Medium confidence and Guess against None were not significant, $t(43) = 1.35$, $p = .19$, and $t(43) = 1.23$, $p = .23$, respectively.

Thus far, analysis suggests that whereas order does not influence the relationship between one judgment type and another, there are not straightforward mappings between confidence and subjective experience judgments. For instance, Remember judgments do

[22] It should be noted that the data presented in Figure 4.1 is not exactly the data that was subject to further analysis as the analysed data was aggregated across First and Second Judgments.

not merely reflect the highest level of confidence, more items were assigned to High confidence than to Remember.

To explore the level of accuracy associated with each response category *a posteriori* probabilities were calculated. Here proportions were based on First Judgments only as for each item both First and Second Judgments were conditional on the same recognition response. Analysis was conducted separately for subjective experience and confidence judgments as the main comparison of interest was accuracy of response categories within judgment type. Missing data was encountered in this analysis as participants had not assigned responses to all response categories; for example, as shown in the *a priori* analysis very few responses were assigned to None or Guess. To increase listwise N these categories were excluded from analysis. As is shown in Figure 4.2, accuracy for confidence judgments decreased as confidence decreased whereas for judgments of subjective experience accuracy was equally high for Remember and Know judgments and only decreased for Familiar and Guess judgments.

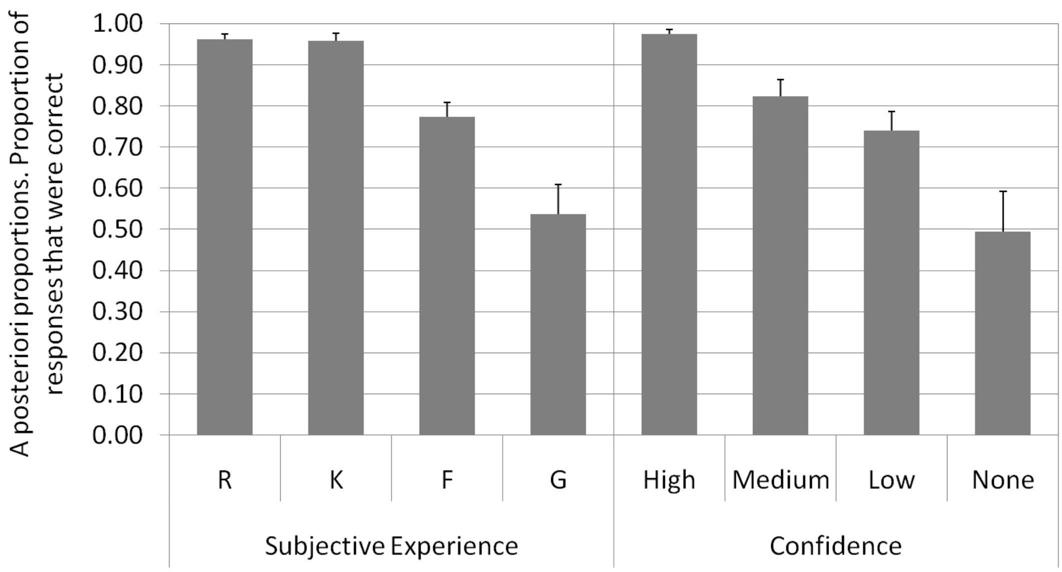

Figure 4.2. Mean proportion of items assigned to Remember (R), Know (K), Familiar (F), and Guess (G) and High, Medium, Low, and None levels of confidence that had previously been correctly recognised. For R, K, and F, N = 40 and for High, Medium, and Low N = 39 due to missing data. Proportions for Guess (N = 24) and None (N = 19) shown but not included in analysis. Errors bars = 1 SeM.

Separate 3(response category) ANOVAs revealed significant main effects of response category for subjective experience judgments, $F(1.60,62.54) = 23.79$, $p < .001$, and confidence judgments, $F(1.57,59.54) = 11.26$, $p < .001$. For subjective experience judgments, no significant difference was found in the accuracy of responses assigned to Remember or Know, $t < 1$, and both Remember and Know responses were significantly more likely to be correct than were responses assigned to Familiar, both $p < .001$. Contrastingly, for confidence judgments, planned comparisons demonstrated that responses assigned to High confidence were significantly more likely to have been correct than responses assigned to Medium confidence, $t(38) = 374$, $p = .001$, or Low confidence, $t(38) = 5.12$, $p < .001$. Medium confidence responses were not found to be more accurate than Low confidence responses, $t(38) = 1.36$, $p = .18$.

In sum, these analyses demonstrate that participants did not use confidence and subjective experience judgments in the same way. The different patterns of distribution of correct responses across confidence levels compared to subjective experience categories demonstrates that participants made significantly more responses as High confidence than Remember. Accuracy analysis further demonstrates the different ways participants used the two judgments. For subjective experience judgments, Know judgments were as accurate as Remember judgments. Whereas High confidence judgments were more likely to be correct than responses assigned to lower levels of confidence. That is, a linear pattern emerges across confidence which is absent for Remember and Know (although a linear pattern is evident for Familiar and Guess). With regards to the order in which the judgments were made, the *a priori* analysis provides initial evidence that the order of the two judgments did not influence distribution patterns across response categories.

4.2.3.3. The influence of one judgment on the other

To fully explore whether making two judgments per item led to one judgment influencing the other judgment, conditional probabilities were calculated on correct recognition decisions using both *a priori* and *a posteriori* methods. Firstly, the proportion of items assigned to Remember, Know, Familiar, and Guess at Second Judgment was calculated split by which confidence level the item had been assigned to at First judgment; i.e., *'of the items assigned to High confidence at First judgment, what proportion were then assigned to each of the categories of subjective experience at Second Judgment?'.* As is shown in Figure 4.3, if an item had been assigned to High confidence at First Judgment then .61 of these items were then assigned to Remember at Second Judgment, .36 were assigned to Know, and .03 were assigned to Familiar. Contrastingly, if an item was assigned to Medium confidence at First Judgment then at Second Judgment .24 were assigned to Remember, .29 to Know, and .47 to Familiar; and so on.

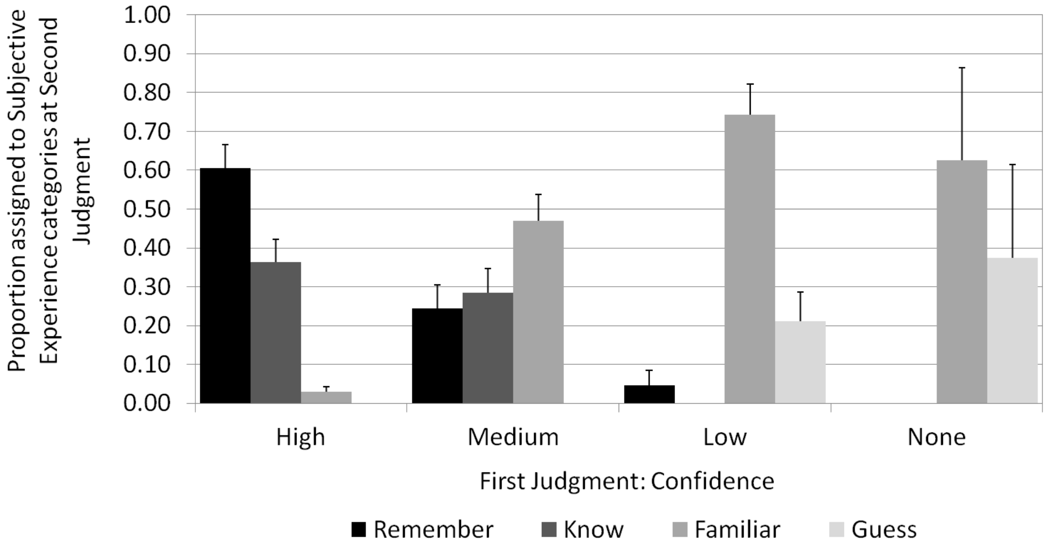

Figure 4.3. Mean proportion of items assigned to each level of confidence at First Judgment that were then assigned to each subjective experience category at Second Judgment. Proportions sum to 1 within confidence level. High N = 43, Medium N = 30, Low N = 26, None N = 4. Data for None shown but not analysed. Error bars = 1 SeM.

These proportions were also calculated in the opposite direction. In the counterbalanced design, participants either judged confidence followed by subjective experience, or the order of these two judgments was reversed. Data was next examined by calculating the proportion of items assigned to Remember, Know, Familiar, and Guess at First Judgment split by which confidence level the item had later been assigned to at Second judgment; i.e., *'of the items that were assigned to High confidence at Second Judgment, what proportion of these had come from a Remember First Judgment?'.* Mean proportions are shown in Figure 4.4.

Comparing Figure 4.3 and Figure 4.4 it is evident that the order in which judgments were made did not influence the relationship between confidence and subjective experience; the graphs are very similar. For example, in Figure 4.4, if an item had been assigned to High confidence at Second Judgment then .64 of these items had previously been assigned to Remember at First Judgment, compared to .61 when the judgments were in the opposing order as shown in Figure 4.3. Across Figure 4.3 and Figure 4.4, within each level of confidence the patterns of means are paralleled for each category of subjective experience.

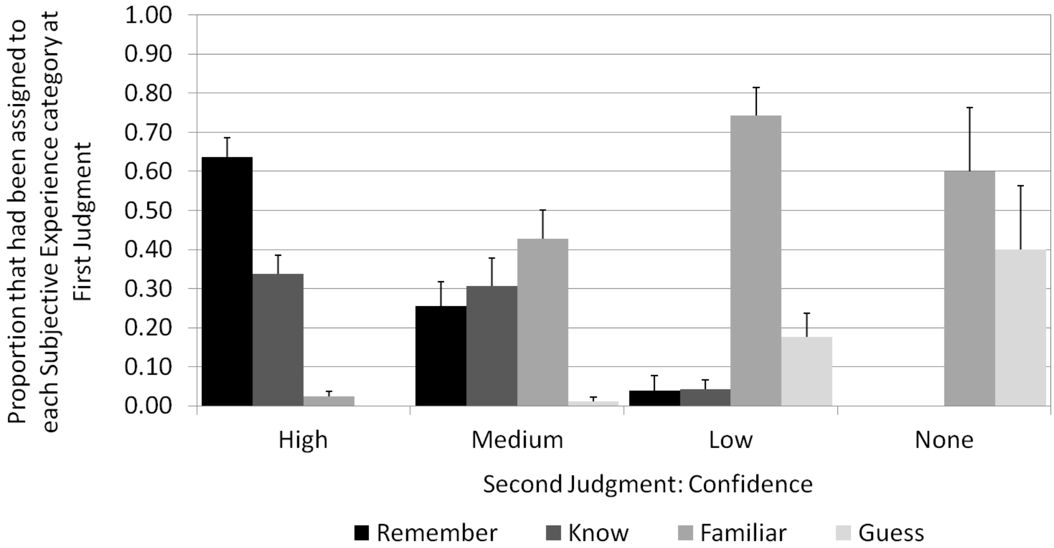

Figure 4.4. Mean proportion of items assigned to each level of confidence at Second Judgment that had previously been assigned to each subjective experience category at First Judgment. Proportions sum to 1 within confidence level. High N = 43, Medium N = 30, Low N = 26, None N = 10. Data for None shown but not analysed. Error bars = 1 SeM.

Separate 2(judgment order) x 4(response category) ANOVAs were conducted for items assigned to each of the confidence levels. These ANOVAs compare the distribution of subjective experience responses at each confidence level across Figure 4.3 and Figure 4.4, i.e., how correct recognition responses that were assigned to High confidence were distributed across subjective experience categories whether judgment order was confidence then subjective experience or the reverse. Effects of judgment order were not calculated as proportions summed to 1.

For High confidence responses the 2(judgment order) x 4(response category) ANOVA demonstrated a significant main effect of response category, $F(1.08,45.37) = 57.87$, $p < .001$, and no interaction between judgment order and response category, $F < 1$. Planned comparisons aggregated across judgment order[23] demonstrated that the proportion of High confidence responses assigned to each level of subjective experience all differed

[23] It should be noted that the means presented in Figure 4.3 and Figure 4.4 were aggregated across judgment order for further analysis which led to an increase in N for some confidence levels. Where t-tests are presented on these aggregated means all results obtained were the same as the result when only listwise N subject to ANOVA were included.

significantly. A greater proportion were assigned to Remember compared to Know, Familiar, or Guess; Know compared to Familiar or Guess; and Familiar compared to Guess; all at least $p < .015$. For Medium confidence responses ANOVA demonstrated a significant main effect of response category, $F(2.00,57.98) = 12.05$, $p < .001$, and no interaction between judgment order and response category, $F < 1$. Planned comparisons aggregated across judgment order demonstrated that for Medium confidence responses a smaller proportion were assigned to Guess than to any other category of subjective experience, all at least $p < .002$, and a greater proportion were assigned to Familiar than to Remember, $t(41) = 2.64$, $p = .012$. There were no significant differences between the proportion of responses assigned to Know and Familiar, $t(41) = 1.37$, $p = .18$, or to Know and Remember, $t < 1$. For Low confidence responses ANOVA demonstrated a significant main effect of response category, $F(1.34,33.42) = 40.38$, $p < .001$, and no interaction between judgment order and response category, $F < 1$. Planned comparisons aggregated across judgment order demonstrated that for Low confidence responses a greater proportion were assigned to Familiar than to any other category of subjective experience, all $p < .001$, and a greater proportion were assigned to Guess than to Remember, $t(37) = 2.75$, $p = .009$, or to Know, $t(37) = 1.34$, $p = .09$, though this only approached significance. There were no significant differences between the proportion of responses assigned to Remember and Know, $t < 1$. ANOVA was not performed on responses with a confidence level of None as the listwise N was 4; less than 5% of responses had been assigned to None. In sum, at each level of confidence there was no interaction between judgment order and subjective experience response, thus demonstrating that the relationship between confidence and subjective experience was not influenced by the order in which the two types of judgment were made.

Conditional proportions were also calculated in the opposing way. Data for proportion of items assigned to each level of confidence was split by the category of subjective experience to which the item had been assigned. Firstly, the proportion of items assigned to High, Medium, Low, and None at Second Judgment was calculated split by which category of subjective experience the item had been assigned to at First judgment; i.e., *'of the items assigned to Familiar at First judgment, what proportion were assigned to each of the confidence levels at Second Judgment?'.* As is shown in Figure 4.5, calculating the probabilities using this method reveals different patterns to those shown in Figure 4.3 and Figure 4.4. If an item was assigned to Remember at First Judgment, .83 went on to be assigned to High confidence at Second Judgment, while .17 went on to be assigned to Medium confidence. Similarly, if an item was assigned to Know at First Judgment, .66 were later assigned to High confidence, while .31 went on to be assigned to Medium confidence. For items assigned to Familiar at First Judgment, .45 and .41 went on to be assigned to Medium and Low respectively at Second Judgment, while for Guess judgments .59 were later assigned to Low and .29 were assigned to None.

These proportions were then calculated in the opposite direction. The proportion of items assigned to High, Medium, Low, and None at First Judgment was calculated split by which category of subjective experience the item had later been assigned to at Second judgment. Mean proportions are shown in Figure 4.6. Comparing Figure 4.5 and Figure 4.6 again demonstrates that the order in which judgments were made did not influence the relationship between confidence and subjective experience. Across Figure 4.5 and Figure 4.6, within each category of subjective experience the patterns of means are paralleled for each level of confidence.

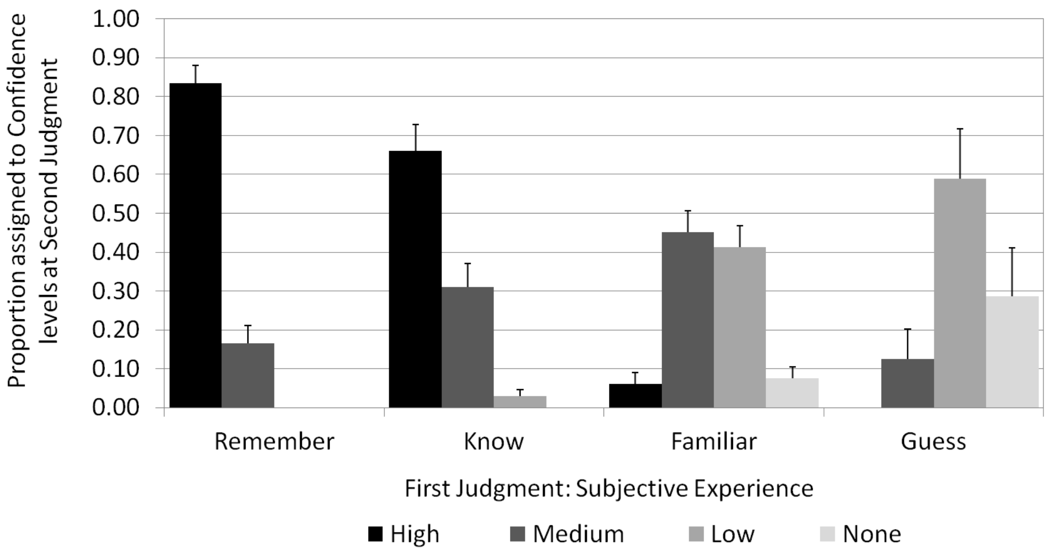

Figure 4.5. Mean proportion of items assigned to each subjective experience category at First Judgment that were then assigned to each level of confidence at Second Judgment. Proportions sum to 1 within subjective experience category. Remember N = 38, Know N = 30, Familiar N = 32, Guess N = 14. Error bars = 1 SeM.

Separate 2(judgment order) x 4(response category) ANOVAs were conducted for items assigned to each of the four categories of subjective experience. These ANOVAs compare the distribution of confidence judgments within each category of subjective experience across Figure 4.5 and Figure 4.6.

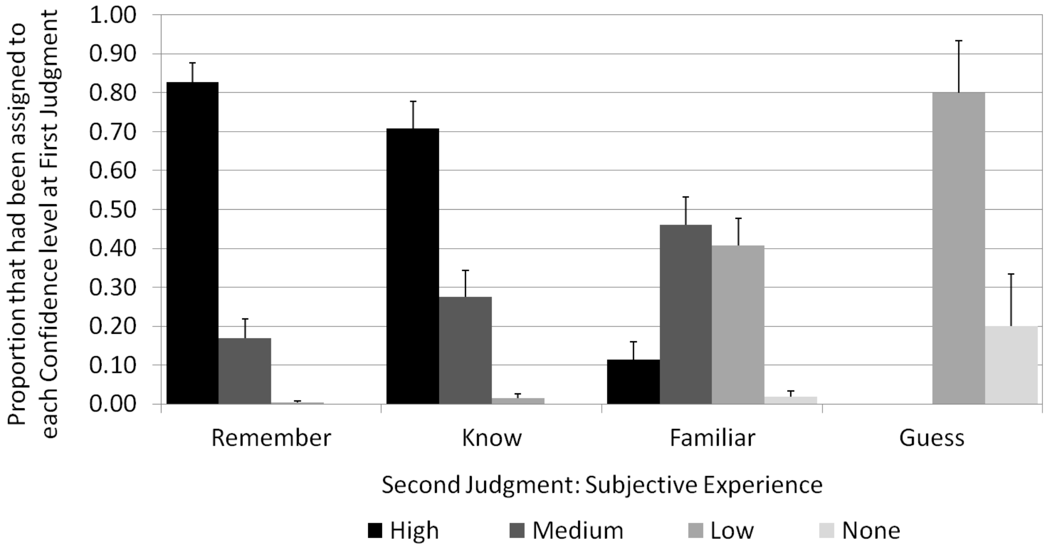

Figure 4.6. Mean proportion of items assigned to each subjective experience category at Second Judgment that had previously been assigned to each level of confidence at First Judgment. Proportions sum to 1 within subjective experience category. Remember N = 38, Know N = 30, Familiar N = 32, Guess N = 10. Error bars = 1 SeM.

For Remember responses the 2(judgment order) x 4(response category) ANOVA demonstrated a significant main effect of response category, $F(1.00,37.09) = 120.38$, $p < .001$, and no interaction between judgment order and response category, $F < 1$. Planned comparisons aggregated across judgment order[24] demonstrated that a greater proportion of Remember responses were assigned to High compared to Medium, Low, or None confidence, and a greater proportion were assigned to Medium confidence compared to Low or None, all at least $p < .01$. There was no significant difference between the proportion of Remember responses assigned to Low and None levels of confidence, $t(42) = 1.47$, $p = .15$. For Know responses ANOVA demonstrated a significant main effect of response category, $F(3,87) = 52.23$, $p < .001$, and no interaction between judgment order and response category, $F < 1$. Planned comparisons aggregated across judgment order demonstrated that all comparisons were significantly different. A greater proportion of Know responses were assigned to High compared to Medium, Low, or None confidence,

[24] It should be noted that the means presented in Figure 4.5 and Figure 4.6 were aggregated across judgment order for further analysis which led to an increase in N for some subjective experience categories. Where t-tests are presented on these aggregated means all results obtained were the same as the result when only listwise N subject to ANOVA were included.

greater proportion were assigned to Medium confidence compared to Low or None, and greater proportion were assigned to Low compared to None levels of confidence, all at least $p < .03$. For Familiar responses ANOVA demonstrated a significant main effect of response category, $F(1.88,58.27) = 20.25$, $p < .001$, and no interaction between judgment order and response category, $F < 1$. Planned comparisons aggregated across judgment order demonstrated that a greater proportion of Familiar responses were assigned to Medium and Low confidence than to High or Low confidence, all $p < .001$. There was no significant difference between the proportion of Familiar responses assigned to Medium or Low confidence, $t < 1$, or between the proportion assigned to High or None levels of confidence, $t < 1$. ANOVA was not performed on Guess responses as the listwise N was 6; less than 5% of responses had been assigned to Guess.

While *a priori* and *a posteriori* methods reveal identical patterns within judgment type and no influence of judgment order, comparing across which judgment type was used to split the analysis demonstrates that participants did not simply map the four levels of confidence onto the four categories of subjective experience, they used the two judgment types in very different ways. Comparing Figure 4.3 to Figure 4.5 we can see that while .61 of items that were initially assigned to High confidence went on to be assigned to Remember, when order of judgments was reversed, .83 of items initially assigned to Remember later went onto be assigned to High confidence. Participants were not 100% confident about an item they Remembered; and if they did have High confidence for an item it did not mean they Remembered it. Comparisons of Familiar subjective experience and Low confidence demonstrate that it was not only Remember and High confidence which did not map onto each other. If an item was judged to be Familiar at First Judgment, .45 and .41 were later assigned to Medium and Low levels of confidence respectively (Figure 4.5). Conversely, if an item was first assigned to Low confidence, .74 then went on to be assigned to Familiar at Second Judgment and .21 went on to Guess (Figure 4.3). Items assigned to Low were often then assigned to Familiar, but not exclusively; and items initially assigned to Familiar were equally likely to be assigned to either Low or Medium levels of confidence.

The above figures are illustrative of the different relationships and influence between successive post-recognition judgments. One further way to analyse this data was to calculate mean confidence ratings for items assigned to Remember, Know, Familiar, and

Guess. This analysis utilised the original numerical confidence levels of 3, 2, 1, and 0 that participants used to make their responses instead of the verbal labels High, Medium, Low, and None. As order of judgment had been shown not to influence patterns of responses above, mean confidence ratings were calculated irrespective of judgment order. Mean confidence ratings for items assigned to each of the subjective experience categories included all confidence judgment regardless of whether that judgment had been made before or after the judgment of subjective experience. As is shown in Figure 4.7, ratings of confidence decreased across the subjective experience categories from Remember, to Know, to Familiar, to Guess, though ratings assigned to Remember and Know items were approximately equal. Due to missing data the Guess category was excluded from the ANOVA. A 3(response category) ANOVA comparing confidence ratings to Remember, Know, and Familiar items demonstrated a significant main effect of response category, $F(3,51) = 73.90$, $p < .001$. Planned comparisons demonstrated that items assigned to Remember or Know were both given higher confidence ratings than items assigned to Familiar: $t(38) = 12.41$, $p < .001$ and $t(38) = 12.73$, $p < .001$ respectively; however there was no significant difference between the confidence ratings given to Remember or Know items, $t < 1$.

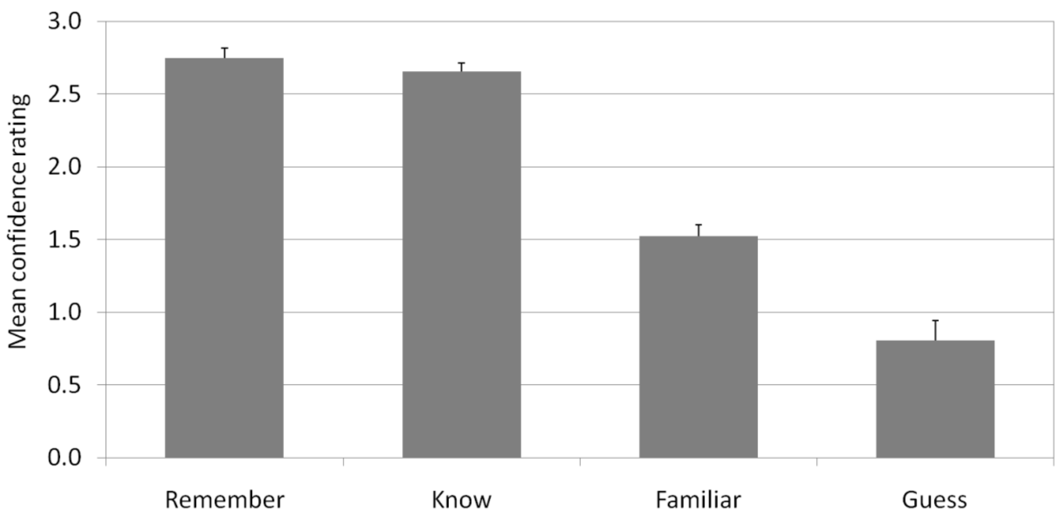

Figure 4.7. Mean confidence ratings given to items assigned to Remember, Know, Familiar, or Guess. N = 39 for Remember, Know, and Familiar. Guess (N = 18) shown but not included in analysis. Error bars = 1 SeM.

In sum, proportional *a priori* and *a posteriori* analysis demonstrated participants did not simply map the four levels of confidence onto the four categories of subjective experience, the four categories of the two judgment types were utilised differently. Analysis of mean confidence examines use of the two judgment types in a different way and demonstrates that items that were assigned to Remember and Know did not differ in terms of confidence. This adds to the previous finding that Remember and Know did not differ in terms of accuracy.

4.2.3.4. Reaction times for recognition decisions

As performed in the experiments presented in Chapter 3, the typical form of analysis to explore reaction time is to analyse RTs for recognition decisions split by which subjective experience category the item was later assigned to. In the current experiment this analysis was performed on correct recognition RTs classified by the confidence or subjective experience response category the item was then assigned to at First Judgment. To increase listwise N the categories of Guess and None were excluded from analysis. Mean correct recognition RTs are shown in Figure 4.8. Both judgment types show similar patterns; recognition decisions based on higher levels of confidence or Remember or Know categories of subjective experience were made faster than recognition decisions based on lower confidence or familiarity.

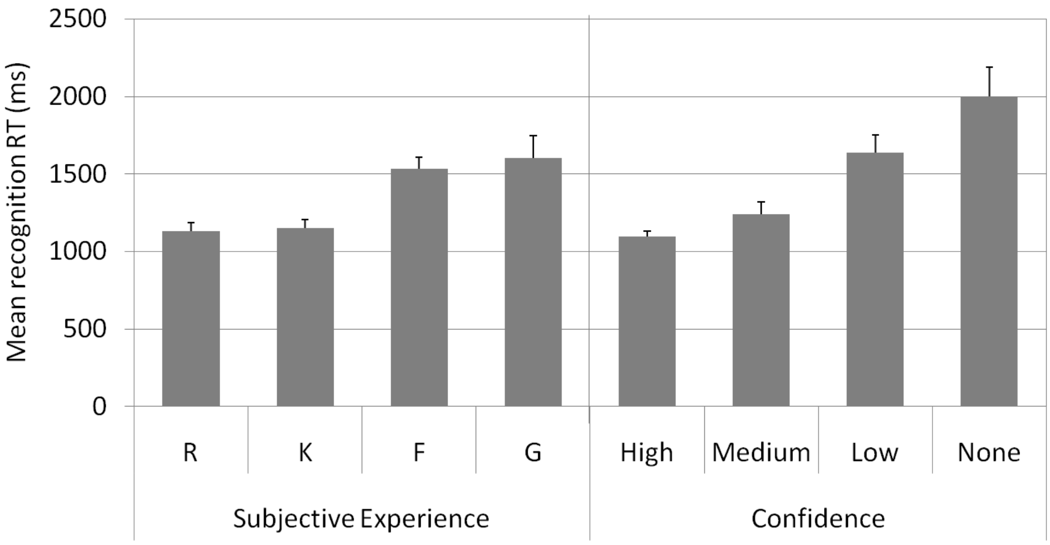

Figure 4.8. Mean correct recognition RTs split by later response category. Means shown are those in the separate subjective experience and confidence 3(response category) ANOVAs. N = 39 for Remember (R), Know (K), and Familiar (F) responses and N = 36 for High, Medium, and Low responses due to missing data. Data for None (N = 12) and Guess (N = 17) shown but not included in analysis. Error bars = 1 SeM.

A 2(judgment type) x 3(response category) ANOVA demonstrated no main effect of judgment type, $F(1,32) = 1.93$, $p = .17$, a significant main effect of response category,

$F(1.38,44.24) = 38.97$, $p < .001$, and no interaction between the two, $F(1.53,48.88) = 1.06$, $p = .35$. Separate 3(response category) ANOVAs for confidence judgments and subjective experience judgments demonstrated significant effects of response category for both: subjective experience, $F(2,76) = 23.04$, $p < .001$, and confidence, $F(1.70,59.66) = 16.42$, $p < .001$. Planned comparisons within judgment type revealed that, for subjective experience judgments, correct recognition decisions to items that went on to be categorised as either Remember and Know were made significantly faster than those for items later categorised as Familiar: $t(38) = 6.05$, $p < .001$ and $t(38) = 5.43$, $p < .001$ respectively. There was no significant difference in the speed with which recognition decisions were made to items that were later assigned to Remember or Know, $t < 1$. For confidence judgments, correct recognition decisions that went on to be assigned to High and Medium confidence were both found to be made significantly faster than decisions where the item was later assigned to Low confidence: $t(35) = 5.08$, $p < .001$ and $t(35) = 3.68$, $p = .001$ respectively. Additionally, recognition decisions made with High confidence were made more quickly than those of Medium confidence, this difference approached significance, $t(35) = 1.89$, $p = .067$. These findings for speed of recognition decision categorised by later subjective experience parallel those in Experiments 3.1 and 3.3 and are the same as those for confidence judgments. Whilst High and Medium confidence responses were not matched on accuracy as Remember and Know responses were, for recognition RTs patterns are matched across the two judgment types.

4.2.3.5. Reaction times for subjective experience and confidence judgments

A novel approach to RT analysis in this experiment was how quickly participants made the actual judgments of confidence and subjective experience. Firstly, the overall mean RTs for confidence and subjective experience judgments to correct recognition decisions were compared irrespective of response category. As shown in Figure 4.9, confidence judgments were made more quickly than judgments of subjective experience. This was confirmed by a 2(judgment type) x 2(judgment order) ANOVA which demonstrated a significant main effect of judgment type, $F(1,43) = 25.51$, $p < .001$. The main effect of judgment order approached significance, $F(1,43) = 3.04$, $p = .09$, as for both judgment types Second Judgments were slightly faster than First Judgments. The interaction between judgment type and judgment order was not significant, $F < 1$.

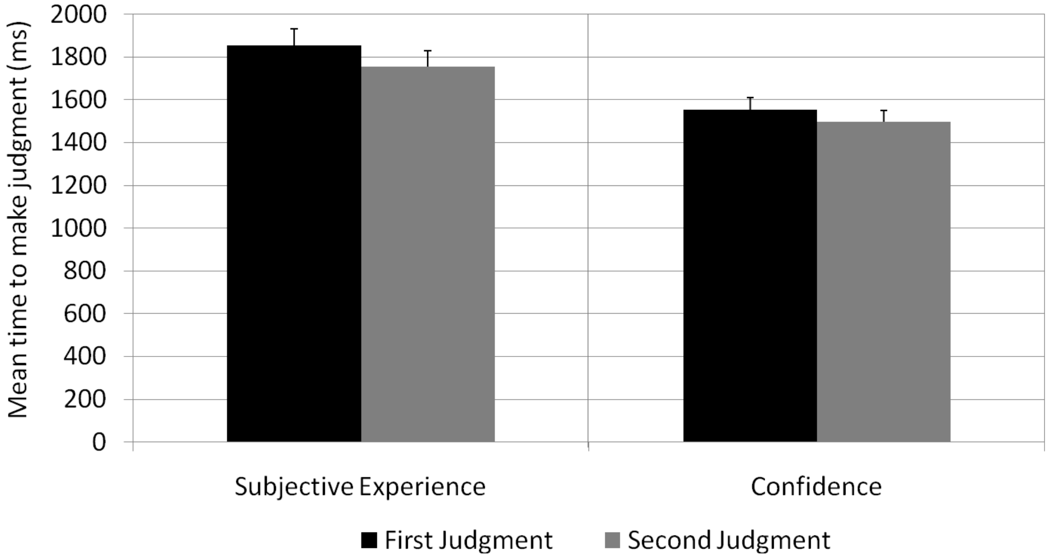

Figure 4.9. Mean judgment RTs (ms) by judgment type and judgment order. Error bars = 1 SeM.

Further analysis was interested in whether differences in RT were observed for different response categories within judgment types. To increase N the categories of Guess and None were excluded from analysis. As the effect of order in the 2x2 ANOVA had not reached significance, data were also aggregated across judgment order. Mean judgment RTs split by response category are shown in Figure 4.10. Separate 3(response category) ANOVAs for confidence and subjective experience revealed a significant main effect of response for confidence judgments, $F(1.25,50.06) = 6.61$, $p = .009$, but no main effect of response for subjective experience, $F < 1$. As can be seen in Figure 4.10, time to make subjective experience judgments was stable at around 2000ms, whilst the time taken to make High, Medium, and Low confidence judgments differed: speed increased as confidence increased.

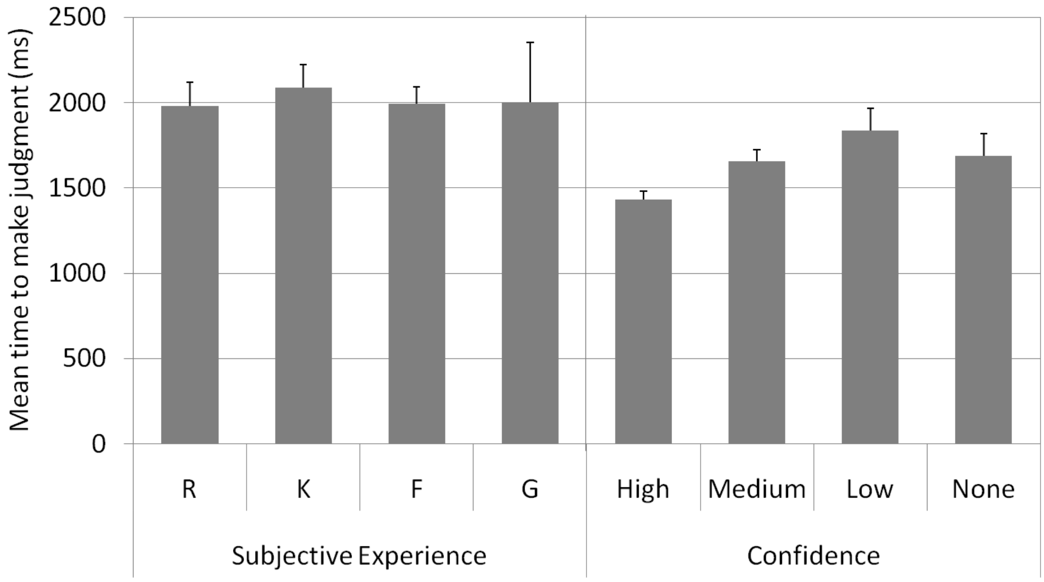

Figure 4.10. Mean judgment RTs (ms) by response category. N = 41 for Remember (R), Know (K), Familiar (F), High, Medium, and Low judgments due to missing data. Data for None (N = 13) and Guess (N = 20) shown but not included in analysis. Error bars = 1 SeM.

Planned comparisons for confidence judgments revealed that High confidence judgments were made significantly faster than both Medium and Low confidence judgments: $t(40) = 4.16$, $p < .001$ and $t(40) = 3.18$, $p = .003$ respectively. The difference in speed between Medium and Low confidence judgments was not found to be significant, $t(40) = 1.34$, $p = .19$. Planned comparisons for subjective experience judgments confirmed that there were no significant differences between the time to make Remember, Know, or Familiar judgments, all at least $p > .26$.

Analysis of judgment RTs demonstrated confidence judgments on the whole were made more quickly than judgments of subjective experience and within these two judgment types contrasting patterns were observed. For subjective experience judgment RTs no effect of response was observed, whilst for confidence judgments the RT advantage observed in recognition judgments for High confidence items was paralleled in the time to make post-recognition judgments of confidence.

4.2.4. Discussion

This experiment compared subjective experience judgments with confidence judgments when both were assessed using four categories of response. The focus of this Discussion is to summarise the central findings of this experiment and more substantive theoretical discussion is carried out in the General Discussion of this Chapter.

Patterns of distributions of responses and accuracy of responses revealed that participants did not use confidence and subjective experience judgments in the same way. Participants were more lenient in their use of confidence judgments than subjective experience judgments and more responses were assigned to High confidence than to Remember. This fits with the findings of Gardiner and Java (1990), Rajaram (1993), Rajaram et al., (2002), and Tunney and Fernie (2007) but demonstrates this without the inclusion of a manipulation such as a word/non-word comparison or masked priming. In the current experiment confidence and subjective experience judgments produced different patterns on identical stimuli.

Accuracy analysis demonstrated a linear relationship between confidence and accuracy, whereas the relationship was not linear for subjective experience judgments as Know judgments were found to be just as accurate as Remember judgments; paralleling the findings presented in Chapter 3. Mean confidence ratings also demonstrated no differences between Known and Remembered items, however both accuracy and confidence ratings differentiated Know from Familiar responses. This important issue is returned to in the General Discussion of this chapter.

Exploration of the influence of one judgment type on the other further demonstrated that participants did not simply map the four levels of confidence onto the four categories of subjective experience. Over 80% of Remember responses were also assigned to High confidence, but when proportions were calculated based on confidence response, only approximately 60% of items which were assigned to High confidence were also Remembered. Whilst subjective experience and confidence judgments are no doubt related, concerns such as those voiced by Bruno and Rutherford (2010) regarding the influence of one judgment on the other do not have credence. The two main points to take from this are that participants did not simply assign all Remember items to High confidence, Know items to Medium confidence, and Familiar items to Low confidence and vice versa; however, the patterns obtained for assignment of responses were matched across judgment order. While all items that were given Remember responses at First Judgment were not assigned to High confidence at Second judgment, the patterns obtained in that judgment order matched the patterns of responding when the order of judgments was reversed. Responses at Second Judgment were not influenced by what response had been made at First Judgment.

RTs were measured at two points: the time taken to make an initial recognition decision and the time taken to make a judgment of confidence or subjective experience. Recognition RT analysis demonstrated a case where confidence and subjective experience responses did not differ. Recognition RTs for subjective experience judgments paralleled those in the Experiments presented in Chapter 3. There was no difference in speed of recognition when items were later categorised as Remember or Know, and recognition of these items was faster than items later assigned to Familiar or Guess. Similarly, recognition of items later classed as High or Medium confidence was equally fast and was faster than for items assigned to Low or None levels of confidence. Taken together with the findings regarding accuracy, where a significant difference between High and Medium confidence responses was demonstrated, this again highlights that subjective experience and confidence judgments behave in different ways.

The novel analysis of comparing RT to make judgments independent of recognition RT confirmed the prediction that confidence judgments would be made more quickly than judgments of subjective experience. It is suggested that it is simply easier to assess memory strength than whether or not you can recollect details from the study episode and this is demonstrated via reaction time. No interaction with judgment order was demonstrated however: confidence judgments were not found to be speeded when made following a judgment of recollective experience.

Confidence judgments were found to be made more quickly the more confident they were but, as predicted, no differences in RT were demonstrated for subjective experience responses. It is suggested that this is related to recollection being a threshold process (e.g., Yonelinas & Parks, 2007). When making a subjective experience judgment, if no recollection of the study episode comes to mind, search is terminated and a different subjective experience response is made. The time to make a judgment of any category of subjective experience is therefore related to the time spent assessing recollection. This appears to be quite a long process compared to making a confidence judgment.

So why might judgments of subjective experience take so long? If making judgments which require recalling information from the encoding phase is a slow and effortful process involving retrieval reaching the threshold of activation for recollection (e.g., Yonelinas & Parks, 2007) or mental time travel (e.g., Tulving, 1985), it might be expected that making a judgment of source will take just as long as making a judgment of subjective experience. Experiments 4.2 and 4.3 were designed with the aim of following up the slower reaction times for subjective experience judgments in Experiment 4.1. However, designing these two experiments also enabled the examination of a number of previously un-researched relationships such as the relationship between source accuracy and confidence and source judgments and Remember, Know, Familiar and Guess subjective experiences.

4.3. Experiments 4.2 and 4.3: Source judgments, subjective experience, and confidence

4.3.1. Comparing subjective experience and source judgments

In contrast to confidence and subjective experience judgments which cannot be objectively verified, source memory judgments can be objectively scored as correct or incorrect. Many researchers have therefore combined judgments of source with confidence and/or subjective experience in attempts to tap both subjective and objective measures of recollection. The central theoretical concern that has arisen from research comparing judgments of source and subjective experience is whether accurate source should be able to be retrieved from memory when an item is assigned to Know. Previous research has found mixed results regarding the accuracy of source for Know items. Some studies have demonstrated source accuracy to only be at chance levels for Know items, which reflects the interpretation of Know responses being recognition unaccompanied by any contextual details from encoding (Dewhurst & Hitch, 1999; Dudukovic & Knowlton, 2006; Perfect, Mayes, Downes, & Van Eijk, 1996, except for Experiment 3). However, other studies have found source accuracy to be above chance for items assigned to Know (Conway & Dewhurst, 1995b; Hicks, Marsh, & Ritschel, 2002; Meiser & Bröder, 2002; Meiser & Sattler, 2007; Starns & Hicks, 2005). This suggests that in some cases enough contextual details from the study phase are retrieved at recognition for source to be judged correctly but not enough details are recollected for the item to be assigned to Remember.

In these studies comparing subjective experience and source there were a number of methodological differences which are relevant for the current experiments. Five of the above eight studies used a standard Old-New recognition paradigm with judgments of source and/or subjective experience being made to only items judged Old, whereas three used a one-step procedure by including a ‘New’ category with the first judgment made. Perfect et al. (1996) and Starns and Hicks (2005) asked participants to make a Remember-Know-New judgment followed by a source judgment and Hicks et al. (2002) had participants make a Left source-Right source-New judgment followed by a judgment of subjective experience. These methodological differences do not map onto the differences in findings however; the five studies which found Know items to have above chance source accuracy are not the five that incorporated a separate Old-New judgment. While one-step and two-step procedures have been shown to influence patterns of responding for Remember-Know judgments (cf., Bruno & Rutherford, 2010; Eldridge et al., 2002; Hicks & Marsh, 1999; see General Introduction, Section 1.6.5.1) and source judgments (Dodson & Johnson, 1993; Marsh & Hicks, 1998) individually, results from studies that employed both types of judgment do not show any clear patterns regarding judgment order.

4.3.2. Comparing source and confidence judgments

Confidence and source judgments have not been directly compared in the same way as confidence and subjective experience: *do these judgments produce the same patterns of responding?* (Gardiner & Java, 1990; Rajaram, 1993; Rajaram et al., 2002), or in the same way as source and subjective experience have: *can source for Know judgments be accurate?* (Conway & Dewhurst, 1995b; Hicks et al., 2002; Meiser & Bröder, 2002; Meiser & Sattler, 2007; Perfect et al., 1996; Starns & Hicks, 2005). Instead, studies comparing confidence and source have typically combined the two judgments into a source confidence judgment. For example, in four experiments Yonelinas (1999) compared recognition confidence and source confidence judgments. These were both measured on 6-point scales from ‘*1 = confident new*’ to ‘*6 = confident old*’, and from *‘1 = sure it was spoken by the female voice’,* to *‘6 = sure it was spoken by the male voice’.* Results demonstrated that the higher the confidence or source confidence judgment made the higher the recognition or source accuracy. In addition, the ROC curves were found to differ for recognition confidence and source confidence and the shapes of the curves obtained were found to fit with the prediction that recognition confidence can rely on both familiarity and recollection while source confidence relies primarily on recollection (Yonelinas, 1999).

Researchers such as Rotello et al. (2005) and Slotnick (2010) have used similar approaches, combining source-confidence measures with recognition-confidence ratings (and judgments of subjective experience) and obtained similar findings with recognition-confidence ratings being higher for accurate, and higher confidence, source judgments. However, the shapes of their ROC curves led these authors to suggest different interpretations of recollection and familiarity. Although these findings are relevant to the ongoing debates surrounding recollection and familiarity processes, they are not of direct interest for the experiments in this chapter.

In relation to the subjective experience and confidence judgments made in these two experiments it is predicted that distribution patterns, accuracy, and RT will parallel those observed in Experiment 4.1. Distribution patterns, accuracy, and RT for these two judgment types are also then able to be explored split by whether the accompanying source judgment is correct or incorrect. As an accurate judgment of source requires retrieval of contextual information from the study episode it is predicted that correct source items will be associated with higher levels of confidence and a greater proportion of Remember responses than incorrect source items. In line with the majority of previous research it is also predicted that while source accuracy will be greatest for items assigned to Remember, a large proportion of items assigned to Know will also be accompanied by accurate source judgments (Conway & Dewhurst, 1995b; Hicks et al., 2002; Meiser & Bröder, 2002; Meiser & Sattler, 2007; Starns & Hicks, 2005).

4.3.3. The influence of one judgment on another

As with the influence of confidence and subjective experience judgments on each other, to date no direct manipulation of judgment order for source and subjective experience judgments has been published. As discussed in Section 4.3.1, experiments that have compared source accuracy and subjective experience judgments have employed a variety of different procedures and findings have been mixed. With regard to the order of source and subjective experience judgments when made after Old-New recognition all five studies which used this procedure asked for subjective experience first and then source[25]. Although no published study has manipulated judgment order for source and subjective experience judgments, one Master's thesis from Jason Hicks' lab has examined this manipulation and has been made available online.

Martin (2007, *unpublished thesis*) manipulated order of source and subjective experience judgments between participants. At study, words were presented either on the left or the right of the computer screen and at test half the participants made a Remember-Know-New judgment[26] followed by a Left-Right source judgment, and the other half made a Left-Right-New judgment followed by Remember-Know. Many of the general patterns of results replicated previous findings. Source accuracy to Remember items was higher than source accuracy for Know items; and source accuracy for Know items was found to be higher than chance (Conway & Dewhurst, 1995b; Hicks et al., 2002; Meiser & Bröder, 2002; Meiser & Sattler, 2007; Starns & Hicks, 2005). However, findings regarding the effect of judgment order were mixed. Participants who had made subjective experience judgments first were found to give more Remember responses to hits than participants who made this judgment following a source judgment, though this finding was only marginal. For source accuracy, no main effect of judgment order was observed when source accuracy was analysed collapsed across subjective experience, though puzzlingly, when response category (Remember-Know) was included in analysis a main effect was reported with source accuracy found to be higher when source was the first judgment made. Additionally, the discussion states that the source-first group showed better source memory for Remember items than the subjective experience-first group, however an interaction between judgment order and subjective experience nor any separate analysis for Remember items was reported. In this unpublished thesis the author does not attempt to explain these discrepancies and it cannot be known which of these opposing findings may be an error in analysis, or an error in reporting. While no study has explored the influence of source and confidence judgments on each other, and these very mixed

[25] Though Conway and Dewhurst (1995b) actually asked participants to make separate Remember-Know judgments for their recognition response and then their source judgment.

[26] The categories utilised by Martin (2007) were actually 'Recollect' and 'Familiar' instead of Remember and Know but for consistency of explanation the standard labels are used here.

findings from an unpublished dissertation should be considered with caution, it is of interest that the influence of one judgment on the other has begun to be explored.

The influence of one judgment on the other is again of key interest in both these experiments. While no effects of judgment order were obtained in Experiment 4.1, it is suggested that the objective nature of source judgments could lead them to have more influence over subsequent judgments. Perhaps when a source judgment is made first and one element of source context is retrieved this may influence the participant to make a Remember or High confidence response somewhat more leniently as was demonstrated, albeit only marginally, by Martin (2007). For example, if ‘only’ the required Left/Right source information is retrieved the item may still be assigned to Remember even if recognition is not accompanied by any feelings of mental time travel or recollection of thoughts that came to mind during study which would have normally be required for the item to exceed that individual’s Remember threshold. Additionally, the accuracy of source judgments could be influenced by whether they precede or follow a judgment of confidence or subjective experience as was suggested by some of Martin’s (2007) data. To examine these suggestions patterns of judgment distribution are explored by judgment order.

4.3.4. Reaction time for judgments of source

Recognition has also been demonstrated to be faster for items when source is accurate compared to inaccurate. Using a one-step source procedure where participants categorised items as either ‘Old - mentally imaged at study’, or ‘Old - read backwards at study’, or ‘New’, Kahn, Davachi, and Wagner (2004) found item + source recognition (Old + correct source) decisions were approximately 250ms faster than item only recognition (Old + incorrect source), but only for items which had been mentally imaged at study. Similar results were obtained by Lundstrom, Ingvar, and Petersson (2005) who also used a one-step procedure. However, opposing results have been demonstrated using a two-step procedure. Wilding and Rugg (1996) presented items in male or female voices at study and correct recognition responses were categorised by whether the subsequent judgment of source was correct. No differences in recognition RTs for correct and incorrect source items were found.

In the current two experiments, analysis of RTs to make post-recognition judgments again allows completely novel analysis to be performed: the comparison of how long it takes to make judgments of source, confidence, and subjective experience. As source judgments require assessment of retrieved contextual information it is predicted that RT for accurate source judgments will be approximately equal to those for subjective experience, which were demonstrated to be slower than confidence judgments in Experiment 4.1. Whether inaccurate source judgment RTs or their prior recognition decision RTs will be slower than those for accurate source judgments will also be explored as some previous research on source recognition RTs has demonstrated differences while other research has not (Kahn et al., 2004; Lundstrom et al., 2005; Wilding & Rugg, 1996).

In these two experiments, influence is also able to be explored in the RT analysis by examining whether the subjective experience category or confidence level an item was assigned to differentiated how long it took to make an accurate source judgment. For example, it could be suggested that accurate judgments of source accompanied by High confidence or Remembering may be made more quickly than accurate judgments of source where confidence is not so high or the recognition is based on Familiarity. Final analysis in these experiments focuses on this original method of exploring judgment RTs.

4.3.5. The current experiments

The materials and basic procedures employed in Experiments 4.2 and 4.3 were identical to those in Experiment 4.1: participants studied medium-frequency words and at test made an Old-New recognition judgment followed by two post-recognition judgments. However, in order that retrieval of source could be tested at recognition, words were presented for study on either the left or the right of the screen with an encoding instruction to associate words on the left with one person and words on the right with another person (after Yonelinas, 1999). The two critical judgments made post-recognition were source and subjective experience in Experiment 4.2 and source and confidence in Experiment 4.3. Statistical comparisons across experiments are left until the Discussion (Section 4.3.8). Subjective experience and confidence judgments were made using the same four response options as in Experiment 4.1 and for source judgments participants selected whether they thought the word had been studied on the Left or the Right. The order in which participant made the two critical judgments was counterbalanced across items.

4.3.6. Methods

4.3.6.1. Participants and Designs

Thirty-two undergraduate students (30 female) participated in Experiment 4.2, mean age 18.88 years (range 18 to 21). Thirty-four undergraduate students (32 female) participated in Experiment 4.3, mean age 18.62 years (range 18 to 20). All participants were from the University of Leeds and received Participant Pool Credits for taking part. Participants were tested in groups of between five and thirteen at individual PCs. In Experiment 4.2, data from two participants was excluded from analysis as they did not follow experimental instructions. Both experiments employed within-subjects designs and for the study and test phases all instructions and stimuli were presented, and data collected, using E-Prime version 1.2.

4.3.6.2. Materials and Procedures

The materials and procedures for Experiments 4.2 and 4.3 were identical to that of Experiment 4.1 except changes due to the source manipulation. In both experiments, in the study phase target words were presented individually to either the left or the right of the computer screen (placement randomised). Participants were instructed that as well as trying to remember all the words, they should also try to remember on which side of the screen the word was shown. To improve performance participants were instructed to associate words on the left with one person and words shown on the right with another person (Yonelinas, 1999). Participants were told they could associate the words with anyone, e.g., celebrities/family/friends, and were asked to write the names of their chosen people on their response booklet (to check they followed this instruction correctly). All other aspects of the study and distracter phases were identical to those of Experiment 4.1.

In the test phases of both experiments participants were instructed that for each word recognised as Old they would then be asked to make two judgments from the list of four judgments. The four judgments were then explained on individual screens. The subjective experience (Experiment 4.2 only), confidence (Experiment 4.3 only), AoA, and pleasantness judgments and instructions were identical to those in Experiment 4.1. For source judgments the screen consisted of two boxes labelled ‘Left’ and ‘Right’ with the question ‘Which side of the screen was this word shown on?’ below. These two response boxes were the same size and shape as the four response boxes required for the subjective experience and confidence judgments. For each source judgment the two response boxes were positioned below the re-presented word, equidistant from the centre of the screen. Participants underwent the recognition test in the same manner as in Experiment 4.1. Each word was presented in the centre of the screen with the cues ‘New’ and ‘Old’ presented below and towards the left and right of the screen respectively, reflecting the number keys which participants were required to use to make their Old-New judgment: 1 for New, 2 for Old. When participants indicated that they recognised a word as Old the first judgment screen was shown. In Experiment 4.2 participants were always required to make either a source judgment or a subjective experience judgment first, followed by another type of judgment (of the three not already made). The critical trials were those where participants made a combination of a source judgment and a subjective experience judgment, either source then subjective experience or subjective experience then source. In Experiment 4.3 participants were always required to make either a source judgment or a confidence judgment first, followed by another type of judgment. The critical trials were those where participants made either source then confidence or confidence then source.

4.3.7. Results

Analysis was conducted in the same manner as in Experiment 4.1 with the addition of comparisons split by source accuracy. Memory performance was examined, followed by overall source accuracy and the proportion of correct responses assigned to confidence levels and subjective experience categories using both *a priori* and *a posteriori* methods. The influence of First Judgment on Second Judgment was then explored using *a priori* and *a posteriori* methods. Reaction times for recognition decisions and judgments were analysed by source accuracy, confidence level, or subjective experience category. Prior to analysis, data from any responses which had been made faster than 300ms or slower than 8000ms were excluded from the dataset. Statistical analysis comparing experiments is left until the Discussion (Section 4.3.8).

4.3.7.1. Memory measures

Memory performance measures for Experiments 4.2 and 4.3 are shown in Table 4.1. In both experiments memory performance was higher than in Experiment 4.1. The source encoding manipulation appears to have increased participants’ encoding proficiency leading to higher performance. Participants made FAs to under 10% of lure items and therefore further analysis of FA data was not performed.

Table 4.1. Mean memory measures (and standard deviations) in Experiments 4.2 and 4.3.

Experiment	Hits	FAs	Recognition performance	*d′*
Experiment 4.2	.71 (.16)	.07 (.06)	.65 (.17)	2.18 (.70)
Experiment 4.3	.79 (.12)	.08 (.08)	.71 (.15)	2.39 (.76)

4.3.7.2. Source accuracy

Judgments of whether an item appeared on the left or right of the screen for study were accurate over 80% of the time in both experiments. Mean proportion of source judgments correct in Experiment 4.2 was .85 (SD = .15) at First Judgment and .84 (SD = .13) at Second Judgment. In Experiment 4.3 proportion correct was .86 (SD = .10) at First Judgment and .88 (SD = .12) at Second Judgment. No effect of judgment order was found in either experiment, both $t < 1$. Source accuracy was high and remained stable whether the source judgment preceded or followed a judgment of confidence or subjective experience. In sum, both memory measures and source accuracy demonstrated that performance was strong and comparative across Experiments 4.2 and 4.3.

4.3.7.3. Subjective experience and confidence measures

The next analysis of interest was which levels of confidence or subjective experiences participants had based their recognition decisions on in the two experiments. Firstly *a priori* proportions were calculated – the proportion of correct responses assigned to each of the confidence levels or subjective experience categories. As is shown in Figure 4.11, there were only slight differences between assignment of First and Second Judgments to responses, but patterns of responses across confidence levels (in Experiment 4.3) differed from patterns of responses across subjective experience categories (in Experiment 4.2). In Experiment 4.3 over 70% of responses were assigned to High confidence, while in Experiment 4.2 just over 40% of responses were assigned to Remember. Approximately equal proportions of responses were assigned to Know and Familiar, whilst slightly more responses were assigned to Medium compared to Low confidence. Also, slightly more responses were assigned to Guess than to None, but both were low, at under 10%.

Patterns for subjective experience responses were very similar to those obtained in Experiment 4.1; however, for confidence judgments more responses were assigned to High confidence in Experiment 4.3 than in Experiment 4.1. This difference is examined further in the Discussion.

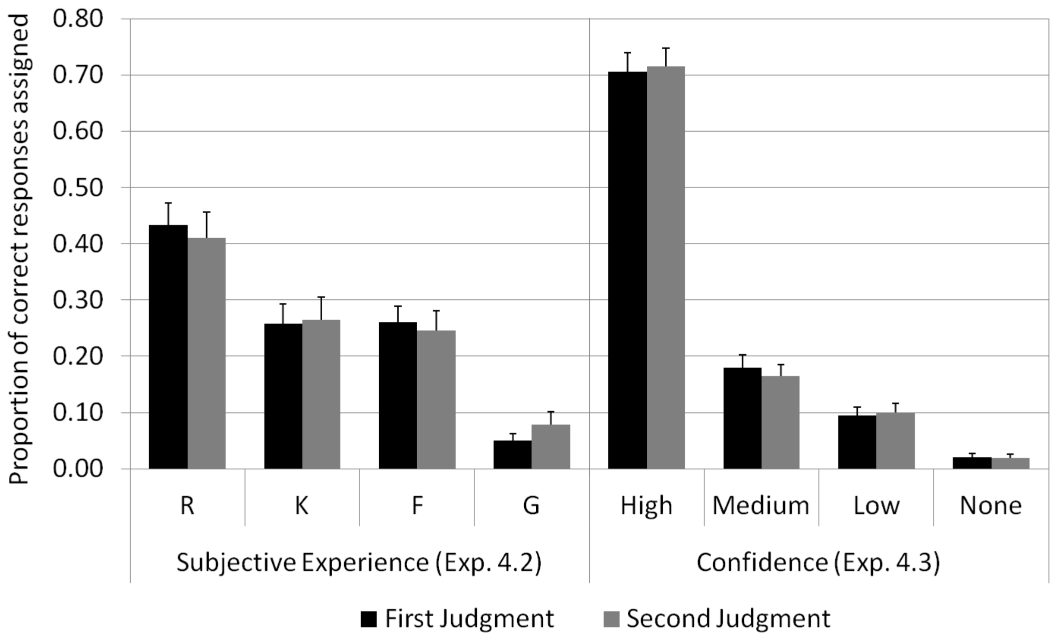

Figure 4.11. Mean proportion of correct responses assigned at First and Second Judgments to subjective experience categories Remember (R), Know (K), Familiar (F), and Guess (G) in Experiment 4.2 and confidence levels High, Medium, Low, and None in Experiment 4.3. Errors bars = 1 SeM.

Separate 2(judgment order) x 4(response category) ANOVAs were performed for Experiments 4.2 and 4.3. For subjective experience responses in Experiment 4.2 there was no main effect of judgment order, $F(1,31) = 1.09$, $p = .31$, a significant main effect of response category, $F(2.06,63.88) = 18.92$, $p < .001$, and no interaction between the two, $F < 1$. Similarly, for confidence judgments in Experiment 4.3 there was no main effect of judgment order, $F(1,33) = 1.34$, $p = .25$, a significant main effect of response category, $F(1.31,43.24) = 236.37$, $p < .001$, and no interaction between judgment order and response category, $F < 1$.

Aggregated across judgment order, planned comparisons were conducted between the different response categories for confidence judgments and judgments of subjective experience[27]. For subjective experience judgments in Experiment 4.2, planned comparisons across categories demonstrated there was no significant difference in the proportion of responses assigned to Know and Familiar, $t < 1$. All other comparisons were significant, all at least $p < .015$; more responses were assigned to Remember than to Know, Familiar or Guess, and fewer responses were assigned to Guess than to any other category. In Experiment 4.3, comparisons across confidence levels demonstrated all proportions were significantly different, all $p < .001$. More responses were assigned to High confidence compared to Medium, Low, or None; more were assigned to Medium compared to Low confidence, and more were assigned to Medium or Low confidence compared to None.

To explore the level of accuracy associated with each response category *a posteriori* probabilities were calculated. Here proportions were based on First Judgments only as for each item both First and Second Judgments were conditional on the same recognition response. Missing data was encountered in this analysis as participants had not assigned responses to all response categories; to increase N the categories of None and Guess were excluded from analysis. As is shown in Figure 4.12, accuracy for confidence judgments decreased as confidence decreased whereas for judgments of subjective experience accuracy for Know judgments was almost as high as for Remember judgments and only decreased for Familiar and Guess judgments.

Separate 3(response category) ANOVAs revealed significant main effects of response category for subjective experience judgments in Experiment 4.2, $F(1.43,31.49) = 20.51$, $p < .001$, and for confidence judgments in Experiment 4.3, $F(1.31,34.23) = 21.40$, $p < .001$. For subjective experience judgments, no significant difference was found in the accuracy of responses assigned to Remember or Know, $t(22) = 1.31$, $p = .20$, and both Remember and Know responses were significantly more likely to be correct than were responses assigned to Familiar, both $p < .001$. In contrast, for confidence judgments, all comparisons were found to be significantly different from each other, all at least $p < .003$. Responses assigned to High confidence were significantly more likely to have been correct than responses assigned to Medium or Low confidence, and Medium confidence responses were found to be more accurate than Low confidence responses.

[27] It should be noted that the data presented in Figure 4.11 is not exactly the data that was subject to further analysis as the analysed data was aggregated across First and Second Judgments.

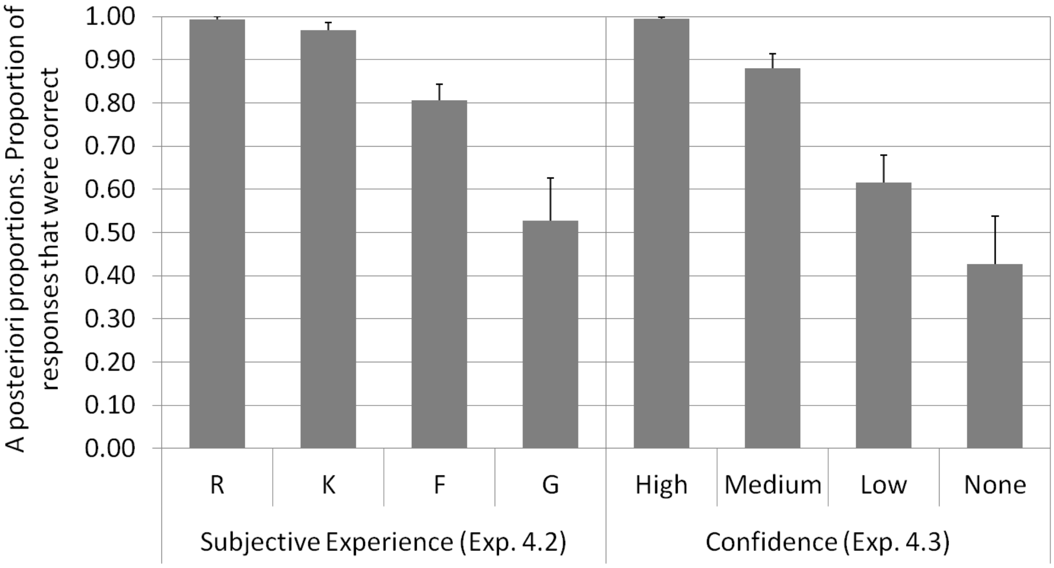

Figure 4.12. Mean proportion of items assigned to Remember (R), Know (K), Familiar (F), and Guess (G) in Experiment 4.2 and High, Medium, Low, and None levels of confidence in Experiment 4.3 that had previously been correctly recognised. For High, Medium, and Low N = 27 and for R, K, and F, N = 23 due to missing data. Proportions for Guess (N = 22) and None (N = 15) shown but not included in analysis. Errors bars = 1 SeM.

Thus, these analyses again demonstrate that participants do not use confidence and subjective experience judgments in the same way. The patterns obtained in these two experiments replicate those obtained in Experiment 4.1. Accuracy analysis again demonstrated that for subjective experience judgments Know judgments were as accurate as Remember judgments, whereas only High confidence judgments were more likely to be correct than responses assigned to lower levels of confidence.

4.3.7.4. The influence of one judgment on the other – Experiment 4.2

To explore whether one judgment influenced the other, conditional probabilities were calculated on correct recognition decisions using both *a priori* and *a posteriori* methods. For Experiment 4.2, firstly the proportion of items assigned to Remember, Know, Familiar, and Guess at Second Judgment was calculated split by whether a First judgment of source had been accurate or inaccurate; i.e., *'of the items where source was judged correctly at First judgment, what proportion were then assigned to each of the categories of subjective experience at Second Judgment?'*. These proportions were then calculated in the opposite direction. In the counterbalanced design participants either judged source followed by subjective experience, or the order of these two judgments was reversed. Data was next examined by calculating the proportion of items assigned to Remember, Know, Familiar, and Guess at First Judgment split by whether the source judgment at Second judgment had been accurate; i.e., *'of the items where source was accurate at Second Judgment, what proportion of these had come from a Remember First Judgment?'. A priori* and *a posteriori* proportions for accurate and inaccurate source judgments are both shown in Figure 4.13[28].

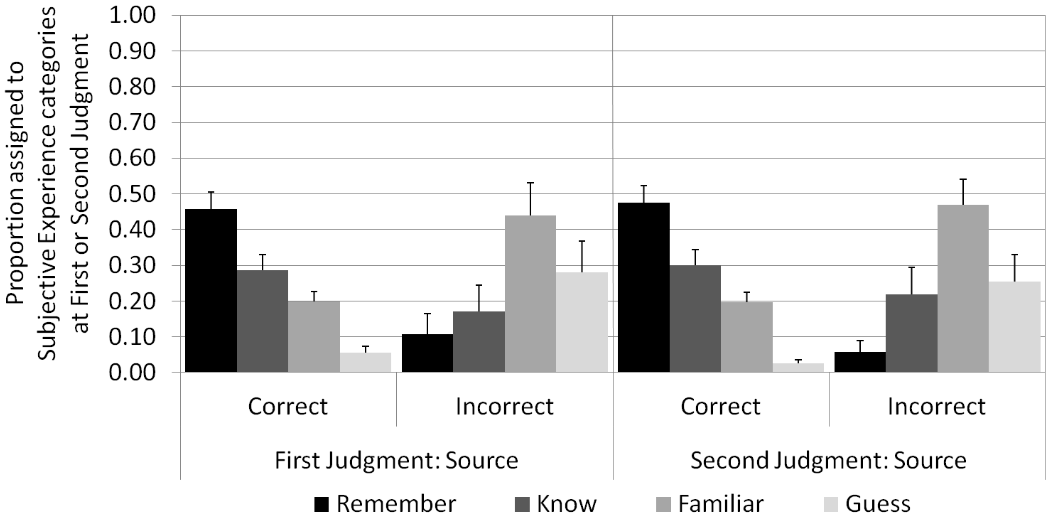

Figure 4.13. The influence of source and subjective experience judgments on one another, Experiment 4.2. Left panel: Mean proportion of items given correct or incorrect source judgments at First Judgment that were then assigned to each category of subjective experience at Second Judgment. Right panel: Mean proportion of items given correct or incorrect source judgments at Second Judgment that had previously been assigned to each category of subjective experience at First Judgment. Proportions sum to 1 within source accuracy and judgment order. Correct Source N = 31, Incorrect Source N = 19 due to missing data. Error bars = 1 SeM.

As is evident from Figure 4.13, opposing patterns were observed when source judgments were accurate compared to inaccurate. For items where source was correct, the majority of these items were associated with Remember judgments. Approximately 45% of items given a correct source judgment were also assigned to Remember, approximately 30% to Know, 20% to Familiar, and less than 5% to Guess. In contrast, for incorrect source judgments (which made up less than 20% of source judgments) the opposite patterns were obtained for Remember, Know, and Familiar judgments. For items where source was incorrect, approximately 45% of items given an incorrect source judgment were assigned

[28] It should be noted that for 'Influence' data in Experiments 4.2 and 4.3 both *a priori* and *a posteriori* proportions are shown together on one figure (per experiment) whereas the comparable data in Experiment 4.1 was separated and shown across two figures.

to Familiar, around 20% to Know, and 6 to 11% were assigned to Remember. For incorrect source items, between 25% and 28% were also assigned to Guess. Separate 2(judgment order) x 4(response category) ANOVAs were performed on correct and incorrect source judgments. Effects of judgment order were not calculated as proportions summed to 1. For correct source judgment items ANOVA demonstrated a significant main effect of response category, $F(1.59,47.73) = 25.72$, $p < .001$, but no interaction between judgment order and response category, $F < 1$. For incorrect source judgment items, ANOVA again revealed a significant main effect of response category, $F(3,54) = 6.34$, $p = .001$, and no interaction between judgment order and response category, $F < 1$.

As judgment order did not interact with response category, further comparisons were conducted with data aggregated across judgment order. For correct source judgment items, planned comparisons demonstrated that a greater proportion of items were assigned to Remember than to any other category of subjective experience, all at least $p < .02$. More correct source items were also assigned to Know than to Familiar and this difference approached significance; $t(30) = 1.96$, $p = .059$. More correct source items were also assigned to Know or Familiar compared to Guess, both $p < .001$. For incorrect source judgment items, planned comparisons demonstrated that a greater proportion of items were assigned to Familiar than to any other category of subjective experience, all at least $p < .02$. No other comparisons demonstrated significant differences, all $p > .25$. Patterns of subjective experience varied systematically depending on whether the source judgment for an item was correct or incorrect. Participants' accurate judgments of source were not just random guesses as the subjective experience accompanying them was appropriate. The greatest proportion of items were assigned to Remember when source was correct whereas the greatest proportion of items were assigned to Familiar when source was incorrect.

These proportions were then calculated in the opposite direction. The proportion of items given an accurate source judgment was calculated split by whether the item was assigned to Remember, Know, Familiar, or Guess. As shown in Figure 4.14, accuracy of source was highest for items assigned to Remember, slightly lower for items assigned to Know, then Familiar, then Guess. Missing data was encountered in this analysis as participants had not assigned responses to all response categories; to increase N Guess was excluded from analysis. A 2(judgment order) x 3(response category) ANOVA was performed on this data.

The main effect of judgment order approached significance, $F(1,19) = 4.12$, $p = .057$, suggesting that perhaps source judgments were more accurate when made as First Judgment (left panel, Figure 4.14). However, overall source accuracy, which took into account all data (Section 4.3.7.2), demonstrated no significant effect of judgment order. There was a significant main effect of response category, $F(2,38) = 15.12$, $p < .001$, and no significant interaction between judgment order and response category, $F(2,38) = 2.20$, $p = .13$.

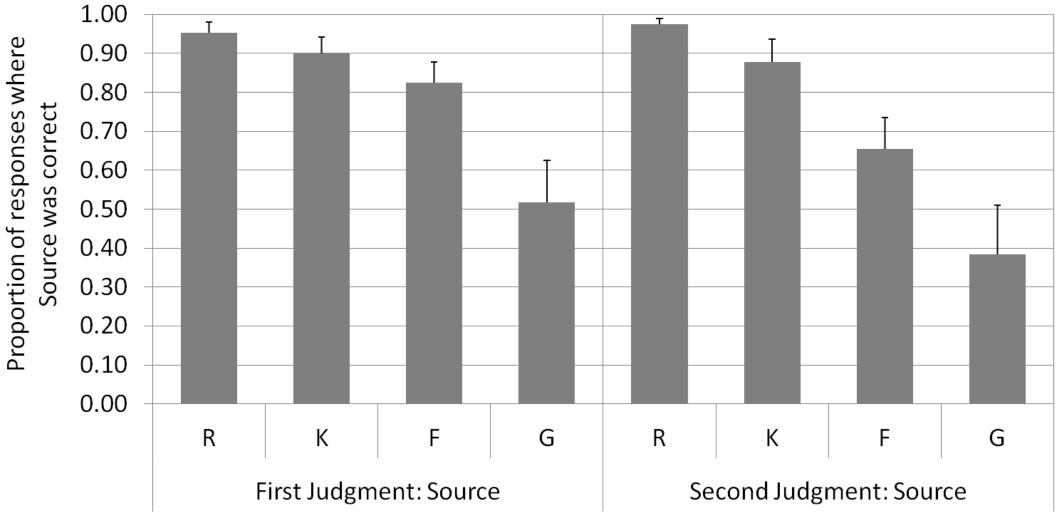

Figure 4.14. Source accuracy split by subjective experience judgment, Experiment 4.2. Left panel: Of the items assigned to each subjective experience category at Second Judgment, the mean proportion given correct source judgments at First Judgment. Right panel: Of items assigned to each subjective experience category at First Judgment, mean proportion given correct source judgments at Second Judgment. N = 20 for Remember (R), Know (K), and Familiar (F) due to missing data[29]. Data for Guess (G) shown but not included in analysis, Ns = 14 and 12 (left and right panels respectively). Error bars = 1 SeM.

Planned comparisons aggregated across judgment order[30] demonstrated significant differences between the accuracy of source judgments to items experienced with different subjective experiences. Remembered items were associated with higher source accuracy than Know or Familiar items, and Know items in turn were associated with higher source accuracy than Familiar items, all at least $p < .02$. One-sample t-tests against chance (.50)

---

[29] Although N is greatly reduced by listwise exclusion for purposes of ANOVA, the inclusive means show the same patterns and are shown in Appendix I.

[30] It should be noted that the data presented in Figure 4.14 is not exactly the data that was subject to further analysis as the analysed data was aggregated across judgment order.

demonstrated that the likelihood of items assigned to Remember, Know, or Familiar being given an accurate source judgment was significantly higher than chance, all $p < .001$. The probability of Guessed items' source judgments being accurate did not differ from chance, $t < 1$. In this experiment where source was a relatively easy 2AFC judgment and deep processing had been undertaken at study, Remember, Know, and Familiar items were associated with above-chance source accuracy.

4.3.7.5. The influence of one judgment on the other – Experiment 4.3

Analysis comparing the influence that judgments of source and confidence had on one another was then performed for Experiment 4.3. Firstly the proportion of items assigned to High, Medium, Low, and None levels of confidence at Second Judgment was calculated split by whether a First judgment of source had been accurate or inaccurate; i.e., *'of the items where source was judged correctly at First judgment, what proportion were then assigned to each of the levels of confidence at Second Judgment?'.* These proportions were then calculated in the opposite direction by calculating the proportion of items assigned to High, Medium, Low, and None levels of confidence at First Judgment split by whether the source judgment at Second judgment had been accurate; i.e., *'of the items where source was accurate at Second Judgment, what proportion of these had come from a High confidence First Judgment?'. A priori* and *a posteriori* proportions for accurate and inaccurate source judgments are shown in Figure 4.15. As is evident from this figure, very different patterns were observed when source judgments were accurate compared to inaccurate. For correct source items, a High level of confidence was generally associated with an accurate judgment of source. Over 75% of items given a correct source judgment were also assigned to High confidence. In contrast, for incorrect source judgments (which made up less than 20% of source judgments) patterns are very different. Approximately 30% of items given incorrect source judgments at first judgment then went on to be assigned to each of High, Medium, and Low levels of confidence at second judgment whereas when judgment order was reversed and source followed confidence, 41% were assigned to Medium confidence and 17% to Low confidence.

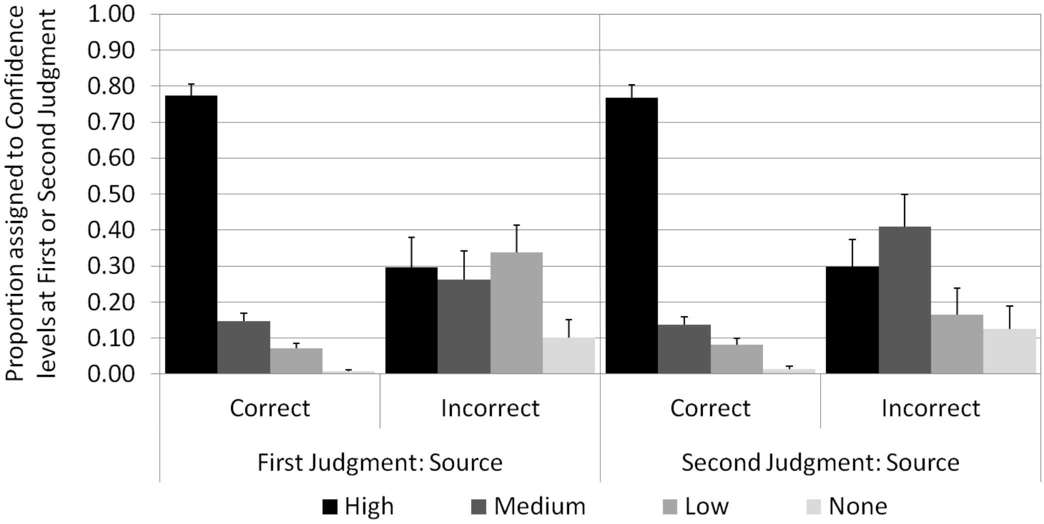

Figure 4.15. The influence of source and confidence judgments on one another, Experiment 4.3. Left panel: Mean proportion of items given correct or incorrect source judgments at First Judgment that were then assigned to each confidence level at Second Judgment. Right panel: Mean proportion of items given correct or incorrect source judgments at Second Judgment that had previously been assigned to each confidence level at First Judgment. Proportions sum to 1 within source accuracy and judgment order. Correct Source N = 34, Incorrect Source N = 23. Error bars = 1 SeM.

Separate 2(judgment order) x 4(response category) ANOVAs were performed on correct and incorrect source judgments. Effects of judgment order were not calculated as proportions summed to 1. For correct source judgments ANOVA demonstrated a significant main effect of response category, $F(3,99) = 334.63$, $p < .001$, but no interaction between judgment order and response category, $F < 1$. As judgment order had not been found to interact with response category, further comparisons were conducted with data aggregated across judgment order[31]. Planned comparisons for correct source judgment items demonstrated that all comparisons between confidence levels demonstrated a significant difference of $p < .001$. When source judgment was correct, a greater proportion of items were assigned to High confidence than to any other levels of confidence; in turn more were assigned to Medium than to Low, and Low compared to None. For incorrect source judgments there was no significant main effect of response category, $F(3,66) = 1.99$, $p = .12$, and no interaction between judgment order and response category, $F(3,66) = 1.63$, $p = .19$. Planned comparisons for incorrect source items demonstrated that there was no

[31] It should be noted that the data presented in Figure 4.15 is not exactly the data that was subject to further analysis as the analysed data was aggregated across judgment order.

difference between the proportion of incorrect source items assigned to High, Medium, or Low confidence, all *t* < 1. However, more incorrect source items were assigned to High, Medium, or Low than were assigned to None, at least *p* < .015. In sum, whilst accurate source items were primarily retrieved with the highest level of confidence, for the less than 20% of items that received incorrect source judgments the confidence judgments that accompanied these items did not vary systematically.

These proportions were then calculated in the opposite direction. The proportion of items given an accurate source judgment was calculated split by whether the item was assigned to a confidence level of High, Medium, Low or None. As shown in Figure 4.16, patterns of source accuracy split by confidence level differed depending on judgment order, with a linear reduction in source accuracy as confidence level decreased when confidence judgments were made following source judgments (left panel), but a more mixed pattern shown when order of judgments was reversed (right panel). Accuracy of source was always highest for items assigned to High confidence and lowest for items assigned to None but source accuracy associated with Medium and Low levels of confidence was variable.

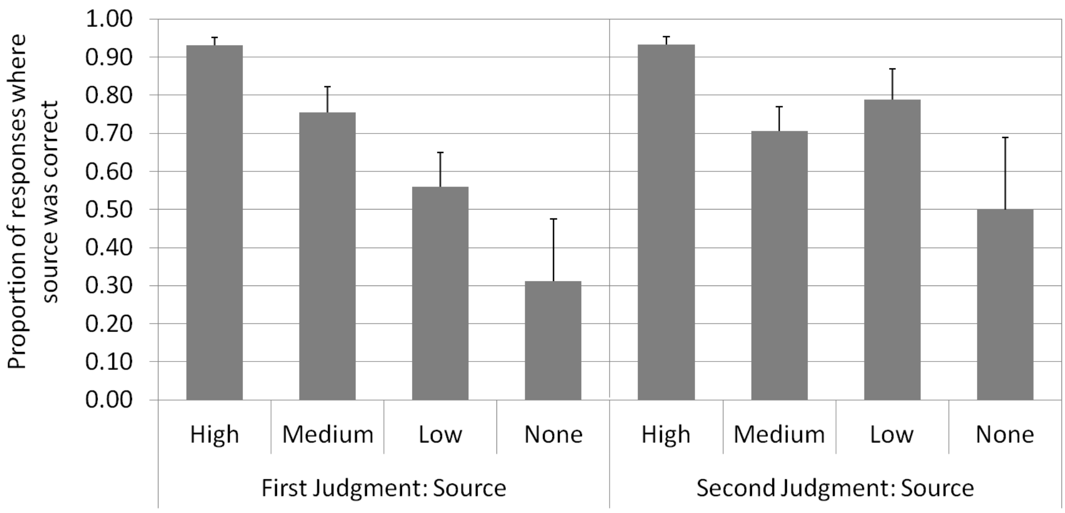

Figure 4.16. Source accuracy split by confidence judgment, Experiment 4.3. Left panel: Of the items assigned to each confidence level at Second Judgment, the mean proportion given correct source judgments at First Judgment. Right panel: Of items assigned to each confidence level at First Judgment, mean proportion given correct source judgments at Second Judgment. Left panel Ns: High = 34, Medium = 30, Low = 24, None = 8. Right panel Ns: High = 34, Medium = 27, Low = 18, None = 7. Error bars = 1 SeM.

A lot of missing data was encountered in this analysis as participants had not assigned responses to all confidence levels. As demonstrated by the distribution patterns of confidence judgments (Section 4.3.7.3) over 70% of correctly recognised items were assigned to High confidence. A 2(judgment order) x 3(response category) ANOVA on this data led to an N of only 13 due to listwise exclusion. Due to the large reduction in N the conditional probabilities in this data were not analysed. Instead, the mean confidence value assigned to correctly recognised items was calculated split by the accuracy of the source judgment assigned to the item. This analysis used the original numerical confidence levels of 3, 2, 1, and 0 that participants used to make their responses instead of the verbal labels High, Medium, Low, and None. Mean confidence for items given a correct source judgment was 2.63 (SD = .28) at First Judgment and 2.58 (SD = .28) at Second Judgment. Contrastingly, mean confidence for incorrect source items was 1.76 (SD = .82) at First Judgment and 1.88 (SD = .83) at Second Judgment. This increased confidence for items that had been given a correct source judgment was demonstrated by a 2(judgment order) x 2(source accuracy) ANOVA which revealed no main effect of judgment order, $F < 1$, a significant main effect of source accuracy, $F(1,22) = 32.98$, $p < .001$, and no interaction between the two. Confidence judgments were significantly higher for items where source judgment was correct compared to items for which source judgment was incorrect.

In contrast to the patterns obtained for subjective experience judgments in Experiment 4.2, confidence judgments varied systematically only for items where judgment was correct; for incorrect source items confidence judgments were evenly spread across responses. Analysis of mean confidence demonstrated that confidence judgments were significantly lower when source was incorrect.

4.3.7.6. Reaction times for recognition decisions

In the current experiments correct recognition RTs were classified by source accuracy at First Judgment or the confidence or subjective experience response category the item was assigned to at First Judgment. To increase listwise N the categories of Guess and None were excluded from analysis in their respective experiments. Mean correct recognition RTs for Experiment 4.2 are shown in Figure 4.17. A 2(response accuracy) ANOVA for recognition RTs categorised by later source accuracy demonstrated that the difference in RTs only approached significance, $F(1,26) = 2.55$, $p = .12$. Correct recognition decisions that were followed by correct source judgments were not made significantly faster than those followed by incorrect source judgments. However, patterns of RTs for judgments of subjective experience replicated those obtained in Experiment 4.1 and a 3(response category) ANOVA for recognition RTs categorised by later subjective experience demonstrated a significant main effect of response category, $F(1.37,30.10) = 4.90$, $p = .025$.

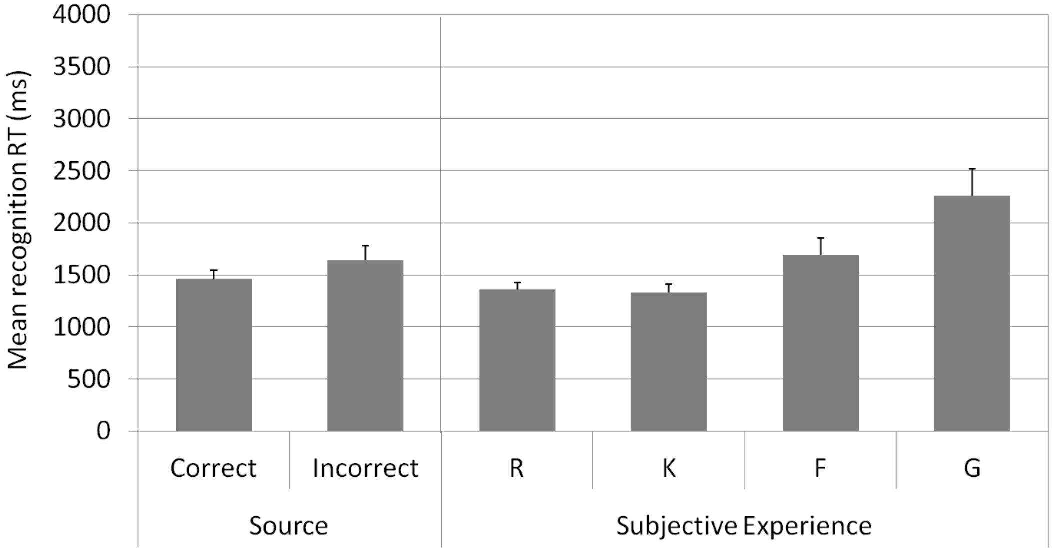

Figure 4.17. Mean correct recognition RTs split by later response category in Experiment 4.2. N = 23 for Remember (R), Know (K), and Familiar (F) responses and N = 27 for source judgments due to missing data. Data for Guess shown but not included in analysis, N = 17. Error bars = 1 SeM.

Planned comparisons demonstrated that correct recognition decisions for items that went on to be categorised as either Remember or Know were made significantly faster than decisions for items later categorised as Familiar: $t(22) = 2.40$, $p = .025$ and $t(22) = 2.31$, $p = .031$ respectively. There was no difference in the speed with which items were correctly recognised if later categorised as Remember or Know, $t < 1$. These findings for speed of recognition decision categorised by later subjective experience parallel those in Experiments 3.1, 3.3, and 4.1.

Mean correct recognition RTs for Experiment 4.3 are shown in Figure 4.18. In this experiment the 2(response accuracy) ANOVA for recognition RTs categorised by later source accuracy found that the difference in means was significant, $F(1,29) = 5.73$, $p = .023$. Correct recognition decisions that were followed by correct source judgments were made significantly more quickly than those followed by incorrect source judgments.

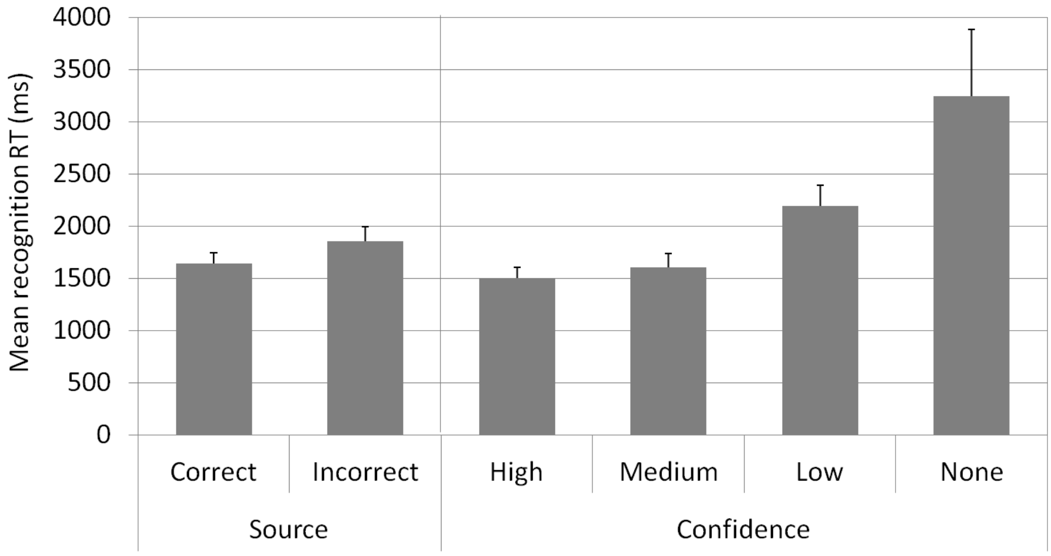

Figure 4.18. Mean correct recognition RTs split by later response category in Experiment 4.3. N = 24 for High, Medium, and Low responses and N = 30 for source judgments due to missing data. Data for Guess shown but not included in analysis, N = 9. Error bars = 1 SeM.

Obviously, Experiment 4.2 and 4.3 do not produce the same result in terms of RT for recognition split by source. As there seemed to be no influence of subjective judgment on source in Experiment 4.2, one might assume that the lack of an effect in Experiment 4.2 was due to a lack of power. Since Experiments 4.2 and 4.3 are identical in terms of memory performance, source accuracy, stimuli, and all elements of methodology (except other critical post-recognition judgment), it is possible to increase power for this analysis by combining the tasks. Across experiments, overall mean correct recognition RT for later correct source items was 1557ms (SD = 511ms) and for incorrect source items was 1753ms (SD = 735ms). Furthermore, a 2(response accuracy) ANOVA demonstrated that, aggregated across experiment, correct recognition decisions that were followed by correct source judgments were made significantly faster than those followed by incorrect source judgments, $F(1,56) = 7.88$, $p = .007$. Thus it could tentatively be concluded that RTs are faster for recognition decisions where source judgments are ultimately correct.

For recognition RTs categorised by later confidence judgment in Experiment 4.3 a 3(response category) ANOVA demonstrated a significant main effect of response category, $F(1.49,34.28) = 10.95$, $p = .001$. Planned comparisons demonstrated that correct recognition decisions for items that went on to be categorised as either High or Medium confidence were made significantly faster than decisions for items later categorised as Low confidence: $t(23) = 3.86$, $p = .001$ and $t(23) = 3.21$, $p = .004$ respectively. There was no difference in the speed with which items were correctly recognised if later categorised as

High or Medium confidence, $t(23) = 1.04$, $p = .31$. These findings for speed of recognition decision categorised by later confidence level parallel those in Experiment 4.1.

4.3.7.7. Reaction times for subjective experience, confidence, and source judgments

These two experiments were again interested in how quickly participants had made their post-recognition judgments. Firstly, time to make source judgments was analysed. Figure 4.19 shows mean source judgment RTs split by accuracy and judgment order for both experiments. A 2(experiment) x 2(source accuracy) x 2(judgment order) ANOVA was conducted with experiment as a between-subjects factor and source accuracy and judgment order as within-subjects factors. Experiment was included as a factor in order to examine whether RTs for source were consistent across experiments or whether source was speeded when participants were also required to report subjective experience or confidence. No main effect of experiment was demonstrated, $F < 1$. ANOVA did reveal a significant main effect of source accuracy, $F(1,46) = 26.49$, $p < .001$, but there was no main effect of judgment order, $F < 1$, and no significant interactions, all at least $p > .30$.

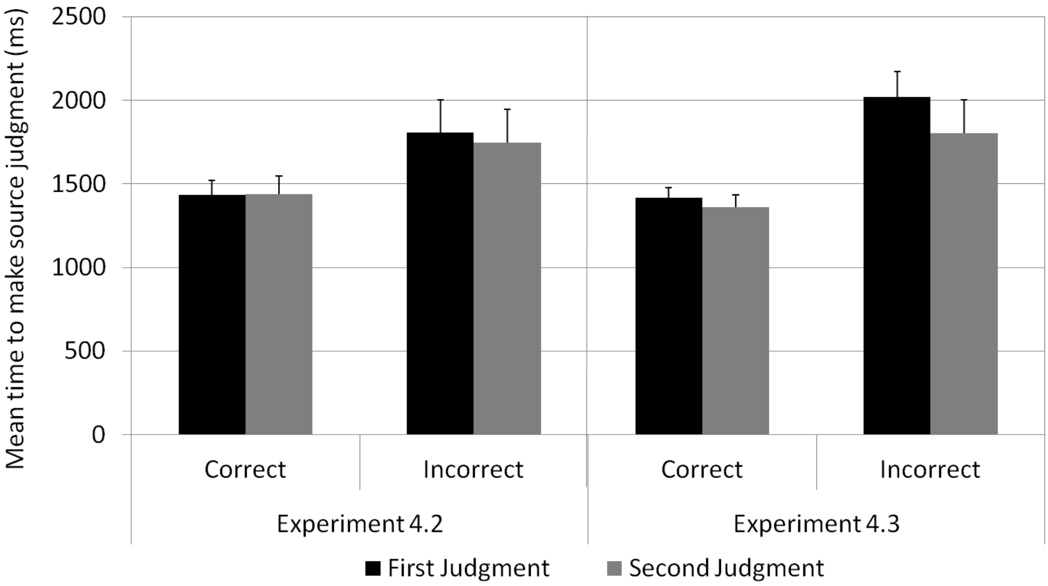

Figure 4.19. Mean source judgment RTs (ms) by accuracy and judgment order in Experiments 4.2 and 4.3. Ns = 24 in both experiments. Error bars = 1 SeM.

Separate planned comparisons for each experiment, with data aggregated across judgment order, confirmed that correct source judgments were made more quickly than incorrect source judgments in both experiments, both $p < .002$. In sum, there were no differences in RT to make source judgments whether they were made as a First or Second judgment; and correct source judgments were made more quickly than incorrect source judgments. Taken together with the recognition RT data these results show a clear picture: when accurate source information does not come to mind when an item is presented RTs are slower for both the initial recognition response and subsequent source judgment.

The next analysis of interest was comparisons between source judgment RTs and RTs to make judgments of confidence or subjective experience. These comparisons were performed using RTs for correct source judgments. Figure 4.20 shows mean time to make source, confidence, and subjective experience judgments in the two experiments.

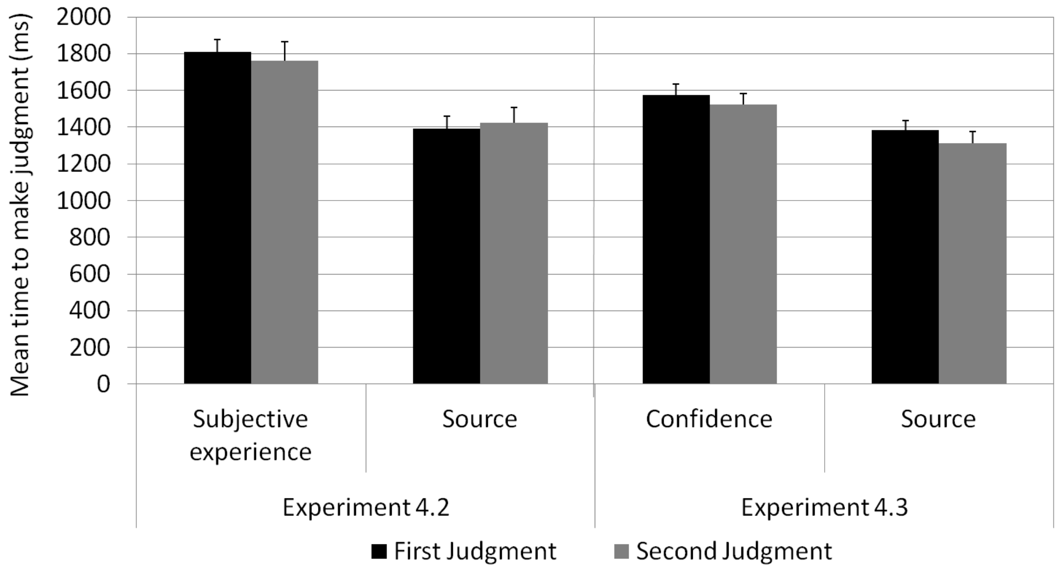

Figure 4.20. Mean judgment RTs (ms) by judgment type and judgment order. Nb. RTs for source judgments here are slightly different to those in Figure 4.19 due to larger N for this analysis. Error bars = 1 SeM.

For Experiment 4.2 a 2(judgment type) x 2(judgment order) ANOVA comparing subjective experience and source RTs demonstrated a significant main effect of judgment type, $F(1,31) = 17.17$, $p < .001$, no main effect of judgment order, $F < 1$, and no interaction between the two, $F < 1$[32]. Source judgments were made significantly faster than judgments of subjective experience and order of judgment did not influence time to make judgment. For Experiment 4.3 a 2(judgment type) x 2(judgment order) ANOVA comparing confidence and source judgment RTs demonstrated a significant main effect of judgment type, $F(1,33)$

[32] An ANOVA using source judgment RT regardless of source accuracy demonstrated the same finding: source judgments were made faster than subjective experience judgments, $F(1,31) = 13.46$, $p = .001$, no main effect judgment order, $F<1$, and no interaction, $F<1$.

= 18.26, p < .001, no main effect of judgment order, $F(1,33)$ = 1.75, p = .20, and no interaction between the two, F < 1[33]. Source judgments were made significantly faster than confidence judgments and order of judgment did not influence time to make judgment. Mean RTs for confidence and subjective experience judgments in these experiments parallel those in Experiment 4.1.

Further analysis was interested in whether differences in RT were observed for different response categories within the confidence and subjective experience judgment types. Guess and None were again excluded from analysis. As the effect of order in the 2x2 ANOVA had not reached significance, data were also aggregated across judgment order. Mean judgment RTs split by response category are shown in Figure 4.21.

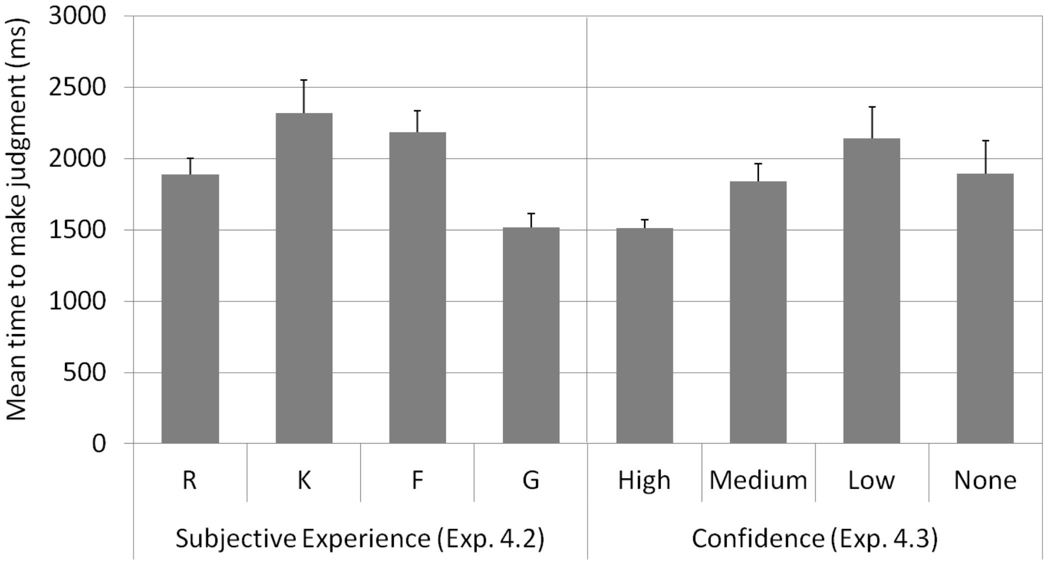

Figure 4.21. Mean judgment RTs (ms) by response category. N = 26 for Remember (R), Know (K), Familiar (F), and N = 28 for High, Medium, and Low judgments due to missing data. Data for None (N = 12) and Guess (N = 20) categories shown but not included in analysis. Error bars = 1 SeM.

Separate 3(response category) ANOVAs for confidence and subjective experience revealed a significant main effect of response for confidence judgments, $F(1.43,38.48)$ = 5.38, p = .016, but no main effect of response for subjective experience, $F(1.50,37.50)$ = 1.84, p =

[33] An ANOVA using source judgment RT regardless of source accuracy demonstrated the same finding: source judgments were made faster than confidence judgments, $F(1,33)$ = 6.32, p = .017, no main effect judgment order, $F(1,33)$ = 2.47, p = .13, and no interaction, F<1.

.18. Results are equivalent to those obtained in Experiment 4.1. As can be seen in Figure 4.21, in Experiment 4.3 time to make High, Medium, and Low confidence judgments differed, with speed increasing as confidence increased. Planned comparisons for confidence judgments revealed that High confidence judgments were made significantly faster than both Medium and Low confidence judgments: $t(27) = 2.83$, $p = .009$ and $t(27) = 2.93$, $p = .007$ respectively. The difference in speed between Medium and Low confidence judgments was not found to be significant, $t(27) = 1.33$, $p = .20$. In Experiment 4.2 the time taken to make Remember, Know, and Familiar judgments was longer and planned comparisons demonstrated that there was no difference between speed of Remember and Know judgments, $t(25) = 1.61$, $p = .12$, or Know and Familiar judgments, $t < 1$. The difference between the speed of Remember and Familiar judgments approached significance, $t(25) = 1.96$, $p = .06$.

4.3.7.8. The influence of one judgment on the other's reaction time

The finding that source judgments were made significantly faster than both judgments of subjective experience and judgments of confidence was not predicted. To explore the relationship between source judgments and the other two judgment types further, RTs to make correct source judgments were analysed split by what category of subjective experience or level of confidence the item was also assigned to. For example, *'if an item was given an accurate source judgment and a High confidence rating, then how quickly was the source judgment made?'.* Source judgment RTs were aggregated across whether the judgment was made First or Second and separate 3(response category) ANOVAs were performed for Experiment 4.2 and 4.3. Guess and None categories were excluded from analysis. Mean source judgment RTs split by the subjective experience category or confidence level the item was assigned to are shown in Figure 4.22. For both experiments, RT to make a source judgment differed across subjective experience categories or confidence levels; the fastest source judgments being those where the item was recognised with High confidence or Remembering. In both experiments ANOVA demonstrated a significant effect of response category: Experiment 4.2, $F(1.46,34.98) = 6.05$, $p = .01$, and Experiment 4.3, $F(1.43,31.45) = 6.24$, $p = .01$.

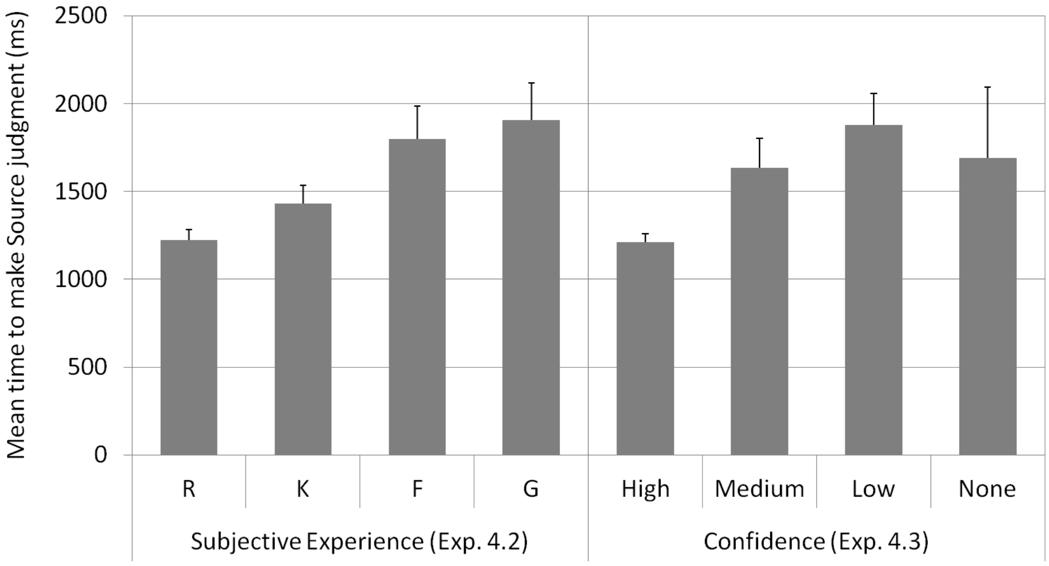

Figure 4.22. Mean time to make correct source judgment split by the subjective experience category (Experiment 4.2, left panel) or confidence level (Experiment 4.3, right panel) the item was assigned to. N = 25 for Remember (R), Know (K), and Familiar (F). N = 23 for High, Medium, and Low. Data for Guess (G) and None shown but not included in analysis, Ns = 14 and 5 respectively. Error bars = 1 SeM.

Planned comparisons in Experiment 4.2 demonstrated that source judgments to Remember items were made more quickly than those for items assigned to Know, $t(24) = 1.94$, $p = .065$, though this only approached significance[34]. Source judgments to both Remember and Know items were also made more quickly than source judgments to Familiar items, Remember: $t(24) = 3.18$, $p = .004$, and Know: $t(24) = 1.84$, $p = .078$, though this only approached significance. Planned comparisons in Experiment 4.3 demonstrated that source judgments to High confidence items were made more quickly than those for Medium, $t(22) = 3.06$, $p = .006$, and Low, $t(22) = 3.73$, $p = .001$, confidence items. There was no significant difference between the time taken to make source judgments for items assigned to Medium or Low confidence, $t(22) = 1.00$, $p = .33$.

These results demonstrate that the subjective experience or confidence associated with recognition of an item was related to how quickly a source judgment was made to that

[34] Based on the literature it would be predicted that Remember judgments are related to judgments of source and therefore a one-tailed t-test could be used here (and the p value obtained would be .033). A one-tailed test was not used as the analysis of source judgment RT split by response category was a post-hoc analysis.

item. This is of particular interest for the relationship between source and subjective experience in Experiment 4.2. In both Experiment 4.1 and 4.2 no clear differences were observed in RTs for Remember, Know, or Familiar judgments; the subjective experience associated with recognition of the item did not influence how long it look for participants to make their judgment of subjective experience (Figure 4.10 and Figure 4.21; although in Experiment 4.2 the difference in speed of Remember and Familiar judgments approached significance). However, the data from Experiment 4.2 concerning source judgment RT data split by subjective experience category demonstrates that, while the type of subjective experience associated with recognition of an item does not influence RTs for judgments of subjective experience, the type of subjective experience associated with recognition of an item does influence RTs for source judgments about the item. This issue is returned to in the Discussion.

4.3.8. Discussion

As predicted, distribution patterns, accuracy, and RT findings demonstrated for confidence and subjective experience judgments in Experiment 4.1 were replicated in Experiments 4.2 and 4.3 demonstrating that participants do not use confidence and subjective experience judgments in the same way. Participants were again more liberal when making confidence judgments compared to subjective experience judgments and accuracy analysis demonstrated that for subjective experience judgments Know judgments were as accurate as Remember judgments, whereas only High confidence judgments were more likely to be correct than responses assigned to lower levels of confidence. Importantly, in these two experiments, patterns of data were also able to be explored split by whether the source judgment made for the item was correct or incorrect.

The source encoding manipulation introduced in these experiments led to differences in memory performance and confidence, but not reports of subjective experience. Recognition performance in both experiments was over 65%, which was higher than the 54% recognition accuracy obtained in Experiment 4.1. Analysis demonstrated that performance in both experiments was significantly higher than performance in Experiment 4.1: Experiment 4.2, $t(74) = 2.49$, $p = .015$, Experiment 4.3, $t(76) = 3.94$, $p < .001$, and across Experiments 4.2 and 4.3 there were no differences between memory performance, $t(64) = 1.34$, $p = .19$, or source accuracy, $t < 1$. Additionally, there were no differences in the proportion of items assigned to Remember, Know, Familiar, and Guess in Experiments 4.1 and 4.2, all $t < 1$. However, in Experiment 4.1 nearly 60% of correct items were assigned to High confidence whereas in Experiment 4.3 this figure was higher, at over 70% (cf. Figure 4.1 and Figure 4.11); significantly more responses were assigned to High confidence in Experiment 4.3 than in Experiment 4.1, $t(76) = 2.71$, $p = .008$ and the corresponding reduction was in the proportion of responses assigned to Low confidence where fewer were assigned in Experiment 4.3 than in Experiment 4.1, $t(76) = 2.94$, $p = .004$. No differences in Medium and Guess proportions were demonstrated across experiments, both at least $p > .11$. The two main points to take away from this are that while the source encoding instructions increased both memory performance and patterns of confidence judgments, they did not alter reports of subjective experience.

Different patterns were also found for confidence and subjective experience when comparing how these judgments were distributed when accompanying source was correct or incorrect. For subjective experience, distribution patterns varied systematically depending on whether the source judgment for an item was correct or incorrect. The greatest proportion of items were assigned to Remember when source was correct whereas the greatest proportion of items were assigned to Familiar when source was incorrect. However, for confidence judgments, whilst accurate source items were primarily retrieved with the highest level of confidence, for items that received incorrect source judgments the confidence accompanying retrieval did not vary systematically. This issue is returned to in the General Discussion (Section 4.4.2).

When proportions were calculated in the opposing direction both Know and Familiar items were found to be associated with above-chance source accuracy in addition to Remembered items. This fits with previous research that has demonstrated that accurate source can accompany retrieval of items assigned to Know (Conway & Dewhurst, 1995b; Hicks et al., 2002; Meiser & Bröder, 2002; Meiser & Sattler, 2007; Starns & Hicks, 2005). While it may appear that the finding of above-chance source accuracy for Familiar items does not fit with this literature, it should be noted that previous studies have conflated Know and Familiar into one judgment. It is interesting that Familiar judgments have been demonstrated to have above-chance source accuracy in the current experiment though, as the majority of results in this thesis have found Know to be similar to Remember (in terms of accuracy and RT, for example) while Familiar is significantly different from Know. Here

Remembered items were found to be associated with higher source accuracy than Know or Familiar items, and Know items in turn were associated with higher source accuracy than Familiar items. Source accuracy was the only factor found to dissociate both Remember from Know and Know from Familiar. For confidence judgments, source accuracy was not able to be calculated for the different levels of confidence as participants had assigned over 70% of items to High confidence. However, as predicted, mean confidence ratings for accurate source items were significantly higher than for inaccurate source items.

Analysis of recognition RTs split by level of confidence or subjective experience at First Judgment replicated the patterns obtained in Experiment 4.1; the patterns observed across all three experiments are considered together in the General Discussion (Section 4.4.3). Analysis of recognition RTs split by source accuracy demonstrated that correct source items were recognised more quickly than incorrect source items. In turn, analysis of judgment RTs also revealed that source judgments were faster when source was correct compared to incorrect and that source judgments were made more quickly than both confidence and subjective experience judgments. This was not predicted as it was thought that RTs for source judgments would be similar to those for subjective experience judgments as both require assessment of what contextual information is able to be retrieved from memory. A possible explanation for this finding is that in these experiments source judgments were able to be made quickly due to the ease of the judgment. Source judgments were only 2AFC (compared to some experimental designs where multiple types of source information are tested at once, e.g., Meiser & Bröder, 2002), and the encoding manipulation whereby participants were instructed to associate words on the left with one person and words on the right with another person led to improved recognition memory and source judgments being correct over 80% of the time. If the task had been more difficult source judgments may require more processing time and RTs may be longer. In future research, source could be made more difficult by increasing the delay between study and test, including multidimensional source judgments, and/or the encoding instructions could be eliminated. It would be of interest to examine which of these manipulations would lead to decreases in source performance while also focusing on whether source RTs and recognition performance are affected. Nevertheless, it is interesting that Experiment 4.2 presents a situation where source is accurate over 80% of the time and RTs for source judgments are fast whereas only 40% of items are assigned to Remember and judgment RTs are slow.

With regard to the influence of one judgment on another, no effects of judgment order were demonstrated and patterns of responses for conditional probabilities were identical whether source had been judged before or after subjective experience or confidence. This does not fit with the findings of Martin (2007) who reported that source judgments were more accurate when performed first and that higher levels of Remember were demonstrated when subjective experience was judged second. Over and above the fact that this data is unpublished and some of his reported findings contradict each other, the procedure used by Martin (2007) conflated the first judgment participants made with the recognition judgment. At test half Martin's participants made a Remember-Know-New judgment followed by a Left-Right source judgment, and the other half made a Left-Right-New judgment followed by Remember-Know. However, one-step and two-step procedures have been demonstrated to produce different patterns of responding for both subjective experience (Bruno & Rutherford, 2010; Eldridge et al., 2002; Hicks & Marsh, 1999) and source judgments (Dodson & Johnson, 1993; Marsh & Hicks, 1998). It is suggested that using a one-step procedure for his first judgment led to the results reported by Martin (2007).

The approach taken in the current experiments was to directly compare source, confidence and subjective experience judgments independently of recognition. Using this procedure the only effect of influence was demonstrated in RTs to make source judgments. For both experiments, RTs to make source judgments differed across subjective experience categories or confidence levels. Source judgments were made more quickly for items recognised with High confidence or Remembering. This demonstrates that the subjective experience or confidence associated with recognition of an item influenced how quickly a source judgment was made to that item. For Experiment 4.3 this finding is in addition to the finding that confidence judgments are made more quickly the higher the confidence. However, the finding is of particular interest for the relationship between source and subjective experience in Experiment 4.2. In both Experiment 4.1 and 4.2 no significant differences were observed in RTs for Remember, Know, or Familiar judgments, the subjective experience associated with recognition of the item did not influence how long it look for participants to make their judgment of subjective experience. It was suggested that this is due to the threshold nature of the recollection process and that memory is assessed until something is recollected or retrieval is terminated and RTs are therefore the same for these two outcomes. However, Experiment 4.2 demonstrates that, while the type of subjective experience associated with recognition of an item does not influence the time it takes to make a judgment of subjective experience, the type of subjective experience associated with recognition of an item does influence the time taken to make a source judgment about the item. This is particularly interesting as source judgments in the current experiments were extremely fast. This finding demonstrates that within that mean time to make an accurate source judgment there was significant variation which was influenced by the subjective experience or confidence with which the item was recognised.

4.4. General discussion

In these three experiments, as no differences in patterns of judgment, relationships with source, or reaction times across judgment order were observed it appears that the fear of contamination of one judgment by the other was unfounded (e.g., Bruno & Rutherford, 2010; Holmes et al., 1998; Rajaram et al., 2002). On the other hand, it could be argued that perhaps all judgments were confounded by the judgment that accompanied them. Comparisons between experiments suggest that this is not the case.

Experiment 4.1 compared subjective experience judgments with confidence judgments in a within-subjects design with judgment order counterbalanced across items. Conversely Experiments 4.2 and 4.3 only employed one of these judgment types and compared it against judgments of source. The patterns for subjective experience and confidence judgments obtained in these two latter experiments are very similar to the patterns obtained in Experiment 4.1, when subjective experience and confidence judgments were made together. These preliminary findings alone strongly support the assertion that making two judgments together does not lead to contamination of one judgment by the other. Comparing of two of the three judgments of interest in each of the experiments allowed both within-subjects comparisons within experiment and between-subjects comparisons across experiments. For example, if subjective experience was influenced when made alongside confidence judgments, then this influence must be identical to that exerted when subjective experience judgments are accompanied by source judgments, as the distribution of subjective experience responses and accuracy of these responses were the same in Experiments 4.1 and 4.2.

4.4.1. Comparing subjective experience and confidence

Overall, patterns of distributions of responses and accuracy of responses revealed that participants did not use confidence and subjective experience judgments in the same way. Participants were more lenient in their use of confidence judgments than subjective experience judgments and more responses were assigned to High confidence than to Remember. This fits with the findings of Gardiner and Java (1990), Rajaram (1993), Rajaram et al., (2002), and Tunney and Fernie (2007). Accuracy analysis demonstrated that Medium confidence judgments were less accurate than High confidence judgments, whereas Know judgments were as accurate as Remember judgments. This latter finding replicates results of Experiments 3.1, 3.2, and 3.3 from Chapter 3 of this thesis. In Experiment 4.1, mean confidence ratings also demonstrated no differences between Known and Remembered items. These findings are of particular note as in Experiments 3.1, 3.2, and 3.3 recognition performance was usually over 80% accurate while in Experiments 4.1, 4.2, and 4.3 performance was 54%, 65%, and 71% respectively. This replication demonstrates the stability of the finding that when you include two separate categories of Know and Familiar, Know responses are just as accurate as Remember responses even with very varied overall recognition performance.

Distribution of responses conditional on previous response and whether the judgment had been made first or second further demonstrated that participants used subjective experience categories and confidence levels in different ways while also revealing that the order in which judgments were made did not influence response patterns. For example, there were no differences in the proportion of High confidence items also assigned to Remember whether the judgment order had been confidence followed by subjective experience or subjective experience followed by confidence. However, there was a difference when comparing the proportion of Remember responses assigned to High confidence with the proportion of High confidence responses that were assigned to Remember. The four categories of the two different type of judgment did not simply map onto each other. Fewer responses were assigned to Remember than to High confidence, and to Low confidence compared to Familiar. In addition, there was no difference in the proportion of responses assigned to Know and Familiar, while there were significant differences between the proportions assigned to all levels of confidence. Subjective experience and confidence are therefore not “experimentally interchangeable” (Rajaram et al., 2002, p. 234), either holistically as judgment types, or when comparing individual categories such as Remember and High confidence.

4.4.2. Comparing source, subjective experience, and confidence

Source judgments cannot be directly compared with subjective experience or confidence judgments in the same way that subjective experience and confidence can be compared as source was a 2AFC judgment with an objectively right or wrong answer; however, the analysis of interest was the relationship between the judgments. Overall, source judgments were accurate over 80% of the time but, as predicted, source accuracy was found to be higher for items recognised with High confidence or Remembering.

Source accuracy analysis was also the only analysis which differentiated Remember and Know responses. Remember and Know were matched on recognition accuracy and recognition RT and while source accuracy was above chance for both Remember and Know responses, the difference between their source accuracy was significant. This suggests that there were differences in how source was retrieved for Remember and Know items. As suggested by Hicks et al. (2002) it appears that judgments of source "can be based on recollection but can also effectively use qualitative characteristics that lack clarity and sufficient amounts of details to give rise to the subjective feeling of remembering" (p. 503). For participants this might mean that retrieval of source is semantic in nature, for example, using the encoding manipulation in the current task a participant might have the thought "I just know I associated *'waltz'* with George Clooney, so it must have been shown on the left, but I can't remember what association I made".

In Experiment 4.2, Familiar judgments were also found to be associated with above-chance source accuracy; though significantly lower source accuracy than for Know items. As discussed previously, although it may appear that this finding does not fit with prior research, those studies did not separate Know and Familiar into two separate categories (Conway & Dewhurst, 1995b; Hicks et al., 2002; Meiser & Bröder, 2002; Meiser & Sattler, 2007; Starns & Hicks, 2005). The current finding suggests that in addition to 'just knowing' the source information in a semantic way, feelings of familiarity at recognition can involve elements of the encoding episode leading to above-chance source accuracy. These findings could be explored further by increasing the difficulty of the source manipulation or increasing the delay between study and encoding. If overall source accuracy is reduced it might be expected that Familiar recognition would no longer have accurate source.

Exploring the relationship between source and other judgments in the opposite direction, patterns of subjective experience and confidence responses split by the accuracy of the accompanying source judgment demonstrated further differences between how subjective experience and confidence judgments were used. For subjective experience responses, a clear relationship between source and subjective experience was demonstrated for both correct and incorrect source judgments. When source was correct a greater proportion of items were assigned to Remember, compared to Know, compared to Familiar. In contrast, when source was incorrect more items were assigned to Familiar than to Know or Remember. Retrieval of source was associated with recollection whilst incorrect retrieval of source was associated with familiarity. Contrastingly, patterns of confidence judgments were only reliable for correct source items. When source was correct, a greater proportion of items were assigned to High, compared to Medium, compared to Low confidence; but when source was incorrect, confidence judgments were evenly split across High, Medium, and Low confidence levels. While this is not incorrect of participants as their confidence judgments were to do with their confidence in their recognition of the item, not their confidence in their source judgment, it shows that there was no relationship between recognition confidence and source when source was incorrect. Critically, it is suggested that judgments of subjective experience are more sensitive and more reflective of memory processes than judgments of confidence. In addition to judgments of confidence being more lenient than judgments of subjective experience, confidence was not found to vary systematically when items were recognised without retrieval of source.

4.4.3. Reaction time for recognition

Recognition was demonstrated to be quicker for items later recognised with Remembering or Knowing, High confidence, or accurate source. These findings fit with the previous literature for confidence (Henson et al., 2000; Mandler & Boeck, 1974; Ratcliff & Murdock, 1976), and source (Kahn et al., 2004; Lundstrom, et al., 2005), and the literature on subjective experience when the split of Know into Know and Familiar is taken into consideration (Dewhurst & Conway, 1994; Dewhurst et al., 1998, 2006; Henson et al., 1999; Rotello & Zeng, 2008; Vilberg & Rugg, 2007; Wixted & Stretch, 2004).

All experiments in this thesis that have measured recognition RTs categorised by later subjective experience response have demonstrated a difference in RT between Remember and Familiar responses, with no differences in RTs for Remember and Know (Experiments 3.1, 3.3 (except 2AFC task), 4.1 and 4.2). Rotello and Zeng (2008) argued against such RT differences being support for dual-process theories by demonstrating that when Remember and Know responses were equated for confidence the RT advantage for Remember responses was reduced. However, the experiments presented in this thesis demonstrate that when Remember and Know responses are equated for accuracy (Experiments 3.1, 3.3, and 4.2) or equated for both accuracy and confidence (Experiment 4.1), both Remember and Know responses are made more quickly than Familiar responses, which had lower levels of accuracy and confidence. This does not support Rotello and Zeng's (2008) conclusion that Remember and Know judgments are based on the same underlying processes and instead demonstrates that splitting Know into separate Know and Familiar categories can reveal important patterns of data otherwise obscured. This key thesis issue is examined further in the General Discussion (Section 5.4).

4.4.4. Reaction time to make judgments

Mean RTs for confidence and subjective experience judgments in Experiments 4.2 and 4.3 paralleled those obtained in Experiment 4.1 and between-subjects comparisons were performed to examine whether there were significant differences in RTs across experiments. No significant differences were demonstrated, all $t < 1$. The time it took participants to make judgments of subjective experience and confidence was equivalent across the three experiments.

In these experiments the novel examination of RTs for source, confidence, and subjective experience judgments, independent of the recognition decision, demonstrated that source judgments were quicker than confidence, which in turn were quicker than subjective experience. This result was not predicted as it was thought that source judgments may take as long as subjective experience judgments as both require assessment of what information can be retrieved regarding context at encoding. However, as discussed already, it is suggested that source judgments may have been so quick in these experiments because it was only a one-dimensional source judgment and encoding proficiency had been improved due to the instructions provided (after Yonelinas, 1999). Future experiments could increase the difficulty of the source judgment or increase the delay between study and test to examine whether these changes influence source judgment RTs. If source judgments are more difficult, or if there is a greater delay between study and test allowing source information to be forgotten, then participants might take longer to make judgments of source.

Examination of source judgment RTs in conjunction with confidence and subjective experience judgments revealed further variation within the already quick source judgments. In addition to source judgments being more accurate when accompanied by Remembering or High confidence, source judgments were also made more quickly when accompanied by Remembering or High confidence. As previously discussed, this is particularly interesting for subjective experience, which itself does not demonstrate a difference in RT for Remember, Know, Familiar, and Guess responses. Critically, the subjective experience associated with recognition of the item does not influence RTs for judgments of subjective experience; however, the subjective experience associated with recognition of the item does influence RTs for judgments of source.

4.4.5. Concluding remarks

The experiments presented here have used novel experimental designs and measures to explore the differences, similarities, and reaction times of confidence, source, and subjective experience judgments. Results demonstrate that making multiple judgments does not influence responding in unforeseen ways (Holmes et al., 1998), however participants use the three types of judgment in different ways, which adds support to the debate that there are differences between confidence and recollection.

5. General Discussion

"The really hard problem of consciousness is the problem of experience*. When we think and perceive, there is a whir of information-processing, but there is also a subjective aspect."* (Chalmers, 1995, p. 201).

5.1. Overview

The central aim of this thesis was to examine how people make and understand judgments of subjective experience. To this end, the experiments presented in Chapters 2, 3 and 4 explored lay conceptions of subjective experience and how these relate to confidence; how participants are able to overcome experimental familiarity; and the relationships between source, confidence, and subjective experience.

Across the ten experiments presented in this thesis the key themes were: How should the subjective experiences of Know and Familiar be defined and understood in recognition memory? What is the relationship between confidence and subjective experience? And how do objective manipulations influence subjective experience? This chapter summarises the central findings of this thesis and presents them in the context of these key themes as well as in the context of broader theoretical discussion surrounding the Remember-Know paradigm, subjective experience, confidence, and familiarity.

5.2. Summary of findings

The main findings from each experiment are shown in Table 5.1 and described in the subsequent sections. Sections 5.2.1 to 5.2.3 discuss the central experimental findings from Chapters 2 to 4 of this thesis and Sections 5.2.4 and 5.2.5 discuss the findings regarding reaction time data across chapters. The implications of these findings for the key themes of this thesis are then discussed in Sections 5.3 through 5.6.

Table 5.1. Main findings from each experiment.

Chapter	Experiment	Main findings
2	2.1	R, K, F, and G justification statements made by other people were rated as significantly more confident from R through to G.
2	2.2	Assignment of High, Medium, and Low confidence K and G justifications to categories of subjective experience was influenced by associated confidence but assignment of R and F justifications was not.
2	2.3	When R, K, F, and G justifications were accompanied by an external confidence value that was appropriate or inappropriate to the justification results of Experiment 2.2 were not replicated; the separate confidence value was not considered in assignment of justifications to categories of subjective experience.
2	2.4	K justifications were able to be reliably split into K and F even though explicit instructions defining the two categories of subjective experience were not provided.
3	3.1	In a between-subjects design, pre-exposure of items impaired recognition but R responses increased. Recollection was required to overcome the familiarity induced by pre-exposure.
3	3.2	In a within-subjects design, pre-exposure impaired recognition but levels of R responses were identical across conditions.
3	3.3	In a blocked design, pre-exposure impaired recognition but levels of R responses were identical across conditions.
4	4.1	The order of confidence and subjective experience judgments did not influence responses. However, the two judgment types were not utilised in the same manner, confidence judgments were more lenient, and K judgments were just as fast and accurate as R judgments.
4	4.2	The order of source and subjective experience judgments did not influence responses. However, whilst RT to make subjective experience judgments did not vary by response, speed of source judgment was faster if accompanied by an R or K response.
4	4.3	The order of confidence and subjective experience judgments did not influence responses. However, RT to make confidence judgments varied by response and speed of source judgment was faster if accompanied by higher confidence.

5.2.1. Chapter 2

Chapter 2 examined how participants interpreted memory justification statements in terms of confidence and subjective experience categories using three internet questionnaires (Experiments 2.1 to 2.3) and one laboratory task (Experiment 2.4). These experiments used the novel methodology of asking participants to judge others' memory justification statements. The justifications used as stimuli in all four experiments were statements made by participants of Gardiner et al. (1997, 1998) to justify their recognition decisions. In Experiment 2.1, participants gave confidence ratings to Remember, Know, Familiar, and Guess justification statements and confidence was found to decrease linearly across those categories of subjective experience. This demonstrated that, in addition to there being a relationship between confidence and subjective experience when these kinds of judgments are a first-person interpretation of the contents of one's own memory (Gardiner & Java, 1990; Dewhurst et al., 2009; Rajaram, 1993; Rajaram, et al., 2002; Rotello & Zeng, 2008; Tulving, 1985; and Yonelinas, 2001), the same relationship exists in how lay people understand others' experiential memory reports. People are experienced in appreciating the relationship between subjective experience and confidence when interpreting their own memory experiences and thus are also able to reliably interpret the memory experiences of others.

The same conclusion was reached from the results of Experiment 2.4 where participants sorted Gardiner et al.'s original Know justifications into two types. In this task, no instructions were given about how justifications should be sorted and the subjective experiences of Know and Familiar were not defined until after sorting was finished. Nevertheless, participants were able to identify two different types of statement within the original Know justifications and these two types mapped onto Know and Familiar. Furthermore, analysis of decision criteria responses identified that Familiar justifications were rated as containing more information and details and Know justifications. The implications of these findings for the separation of Know and Familiar are discussed in Section 5.3.

Experiments 2.2 and 2.3 used the confidence ratings obtained in Experiment 2.1 to select justifications for further study and the task for participants was to assign justifications to categories of subjective experience. Experiment 2.2 used high, medium, and low confidence statements from each subjective experience category and demonstrated that when the confidence value associated with a justification differed, Remember and Familiar justifications were consistently categorised to Remember and Familiar respectively, whereas the majority of low confidence Know justifications and high confidence Guess justifications were categorised as Familiar. Through examining the content of the justifications, it was suggested that these differences occurred because Remember and Familiar justifications at any confidence level directly described the processes underlying the recognition decisions whereas the cognitive processes reflected in Know and Guess justifications differed with confidence. A key theme of this thesis is that Know responses are considered to reflect high confidence without recollection and Guess responses are considered to be low-confidence familiarity-based responses. Thus, if Know and Guess responses are related to one's confidence in evaluating the contents of memory then it follows that categorisation of these justifications would be influenced to a greater extent by confidence level when it is manipulated experimentally. Indeed, Geraci et al. (2009) demonstrated that different patterns of responding are observed if confidence is emphasised in how Know is defined to experimental participants. The current experimental methodology was very different from Geraci et al.; however, this comparison demonstrates that the same patterns are observed whether it is first-person subjective experiences under study or third-person interpretations of others' experiential reports.

Experiment 2.3 followed this up by manipulating confidence more overtly. Here prototypical justification statements from each category's mean confidence level were presented accompanied by a confidence value that was either appropriate (to that subjective experience) or inappropriate (from the confidence values of a different subjective experience category). Here no effects of confidence were observed. Remember justifications were consistently categorised as Remember; Familiar were consistently assigned to Familiar; and assignment of Know and Guess justifications was split between their original category and Familiar. Though this might appear to be a negative result – that no influence of confidence was demonstrated – instead it is taken to reflect the implicit understanding of participants that the experiences of memory are more important than confidence and that confidence is derived from subjective experience, not the other way round (Gardiner, 2001; Tulving, 1985). If the confidence value accompanying a justification did not fit with the justification, participants aligned their responses to the experiential state and not the confidence level given.

On the whole, the results of Chapter 2 suggest that confidence is a process which is in some part reliant upon inferences made during retrieval – the fact that appropriate confidence judgements can be invoked using third-person justifications of retrieval suggests that to some extent they may be 'rules of thumb' imposed on information produced at retrieval. Given that confidence level, however, does not seem to influence the category of experience, the reverse cannot be claimed – reading off another's trace strength does not recover information about recollective status.

5.2.2. Chapter 3

In Experiments 3.1, 3.2, and 3.3 the familiarity of some or all items was manipulated prior to study using a pre-exposure vowel-counting task. This manipulation led to impaired recognition performance for pre-exposed items in all experiments, replicating previous findings by Åberg & Nilsson (2001, 2003), Dobbins, Kroll, Yonelinas, et al. (1998), Kormi-Nouri et al. (2005), Chalmers and Humphreys (1998), Greene (1999), Maddox and Estes (1997), and Tulving and Kroll (1995). However, subjective experience responses were only found to differ between exposure conditions when exposure was manipulated between-subjects (Experiment 3.1). When pre-exposure was manipulated using a within-subjects design (Experiment 3.2) or a blocked within-subjects design (Experiment 3.3) patterns of subjective experience responses were the same across exposure conditions demonstrating carry-over from one condition to the other (e.g., Conway et al., 2001).

In all three experiments, when any or all items had been pre-exposed, Remember responses were positively correlated with recognition performance while Familiar responses were negatively correlated with performance. The importance of methodological considerations in subjective experience experiments is returned to in Section 5.6. The current findings fit with the conclusions drawn by Dobbins, Kroll, Yonelinas, et al. (1998) and Whittlesea (2002) – manipulations imposed on stimuli can influence the subjective experiences used as the basis for recognition. Here, depending on whether pre-exposure was manipulated within- or between-subjects, subjective experience responses differed depending on how an item was perceived in relation to the other items on the test (see also Bodner & Lindsay, 2003; Bodner & Richardson-Champion, 2007; and Dewhurst & Parry, 2000).

The experiments in Chapter 3 were the first to manipulate pre-exposure between-subjects and the first that had examined subjective experience responses in pre-exposure. Tulving and Kroll (1995) and Åberg and Nilsson (2003) had had participants make confidence judgments on a 3-2-1 scale (hypothesised to map onto Remember-Know-Guess judgments by Åberg & Nilsson, 2003) and had found higher confidence for correct decisions when items *had not* been pre-exposed; however, the opposite pattern was observed in Experiment 3.1 as more Remember responses were obtained for items that *had* been pre-exposed. As discussed above, although confidence and Remember-Know judgments are no doubt related, this example highlights how the two can produce very different patterns of responses. The relationship between confidence and subjective experience is returned to in discussion of Chapter 4's experiments and theoretical implications are examined in Section 5.4. Results of Chapter 3's experiments were interpreted as supporting the source discriminability interpretation of pre-exposure (Dobbins, Kroll, Yonelinas, et al., 1998) not the novelty-encoding hypothesis (Tulving et al., 1994, 1996; Tulving & Kroll, 1995). Novelty at study did not improve recognition performance for non-pre-exposed items, familiarity impaired recognition performance for pre-exposed items. Critically, the very fact that subjective experience was carried-over from pre-exposed to non-pre-exposed items suggests that the memory impairment in pre-exposure was due to retrieval of items from memory, not from their encoding. Participants were not able to determine which items had been pre-exposed and therefore which they would *need* to use recollection to recognise so they used recollection as the basis for all recognition decisions.

5.2.3. Chapter 4

In Chapter 4, three experiments examined the relationship between judgments of confidence, source, and subjective experience, when two judgments were made following each recognition decision. Previous researchers had suggested that when two judgments were made one judgment might be contaminated by the other judgment, for example, all Remember responses being assigned to High confidence (Bruno & Rutherford, 2010; Holmes et al., 1998; Rajaram et al., 2002). The current experiments found no differences in patterns of judgment, relationships with source, or reaction times across judgment order and patterns were consistent across experiments demonstrating that the fear of contamination was unfounded. Critically however, patterns of responses revealed that relationships between each of the three judgment types are different.

Experiment 4.1 demonstrated that confidence and subjective experience categories are not used in the same way. Patterns of distributions of responses and accuracy of responses revealed that participants were more lenient in their use of confidence judgments than subjective experience judgments. This supports the findings of Gardiner and Java (1990), Rajaram (1993), Rajaram et al., (2002), and Tunney and Fernie (2007), but the current experiments were the first to demonstrate this by contrasting the four categories of Remember, Know, Familiar, and Guess against four levels of confidence. Subjective experience and confidence were demonstrated to not be "experimentally interchangeable" (Rajaram et al., 2002, p. 234), either holistically as judgment types or when comparing individual levels within judgment types. Furthermore, when a source encoding manipulation was introduced (in Experiments 4.2 and 4.3), overall recognition performance increased and use of confidence levels changed but patterns of subjective experience responses were unaltered. This finding is discussed further in Section 5.4.

Comparing source judgments and judgments of subjective experience and confidence in Experiments 4.2 and 4.3 demonstrated that source accuracy was highest for items recognised with High confidence or Remembering, however Know and Familiar responses also achieved above-chance source accuracy[35]. That Know responses were accompanied by accurate source judgments supports the previous findings of Conway and Dewhurst (1995b), Hicks et al. (2002), Meiser and Bröder (2002), Meiser and Sattler (2007), and Starns and Hicks (2005) who found above-chance source accuracy for Know items. It appears that for a significant proportion of items which are Known, enough contextual details from the study phase are retrieved at recognition for source to be judged correctly but not enough details are recollected for the item to be assigned to Remember.

These findings go against the previous research that had not found above-chance source accuracy for Know responses (Dewhurst & Hitch, 1999; Dudukovic & Knowlton, 2006; Perfect et al., 1996, except for Experiment 3); however, methodological differences could explain why these studies demonstrated opposing findings to those discussed above. For example, Dudukovic and Knowlton (2006) only measured retrieval of source details in a re-test which took place one week after initial testing. This delay led to a decrease in Remember responses and an increase in Know responses indicating an overall reduction in retrieval of contextual details. It can be suggested that this forgetting of source information is what led to items being Known rather than Remembered and why Know

[35] Medium and Low confidence responses could not be analysed in terms of source accuracy as 70% of items had been assigned to High confidence resulting in a lot of missing data.

responses were not accompanied by accurate source judgments. Dewhurst and Hitch (1999) did not employ a delay but in their study participants were only required to make a source judgment if they were confident about their recognition decision. In light of this, Dewhurst and Hitch (1999) suggested that the low response rate for source might have led to the low source accuracy for Know items. For Perfect et al. (1996) it is the difficulty of the source judgment that can be suggested to have led to at chance source accuracy for Know items. Of their four source accuracy experiments, the three which found Know responses were not accompanied by accurate source employed a temporal order source judgment (Experiment 1), a List 1/List 2 source judgment where there were no differences between lists (Experiment 2), and a quadrant presentation where items were presented in one of four spatial locations but presentation was sequential (Experiment 4). These source manipulations are more difficult than that employed in their Experiment 3, which did find above-chance source accuracy for Know. Experiment 3 used the quadrant presentation but four stimuli were presented simultaneously. This would have allowed participants to more easily create associations between items. Some of these associations may have incorporated the spatial relationship between items and thus may have aided retrieval of location at test even if other elements of the association were not recollected and the experience of retrieval was one of Knowing rather than Remembering.

That Familiar responses were also accompanied by above-chance source accuracy had not been demonstrated previously; the current experiments were the first to separate Know and Familiar when comparing subjective experience and source judgments. The current finding suggests that feelings of familiarity at recognition can include retrieval of elements of the encoding episode leading to above-chance source accuracy. It is not just recognition accompanied by recollection or recognition based on a subjective experience of high confidence without recollection that can involve retrieval of contextual information from encoding, in the current experiments recognition based on familiarity also had an above-chance likelihood of being accompanied by retrieval of correct source information. As suggested in Chapter 4, future research could examine whether Know and/or Familiar responses lose their accurate source when the source judgment is made more difficult by increasing the delay between study and test, requiring multidimensional source judgments, or eliminating the deep encoding instructions that were used in the present experiments. Indeed, previous research has demonstrated that the specificity of the source judgment differentiates between whether Know responses have accurate source or not.

Meiser and Sattler (2007) found that specific source memory for the voice of an individual speaker was more accurate for Remember items than for Known items. However, partial source memory for the more global attribute of gender did not differentiate between Remember and Know items.

To return to a critical theme of this thesis, further differences between subjective experience and confidence were observed when responses were split by the accuracy of the accompanying source judgment. For subjective experience, retrieval of source was primarily associated with recollection whilst incorrect retrieval of source was mainly associated with familiarity. Contrastingly, for confidence judgments, retrieval of source was primarily associated with High confidence but there was no relationship between recognition confidence and source when source was incorrect. This was taken to suggest that judgments of subjective experience are more sensitive and more reflective of memory processes than judgments of confidence; this issue is returned to in Section 5.4.

5.2.4. Reaction time for recognition decisions

Five experiments in this thesis examined reaction times to make recognition decisions. This analysis compared RTs for correct recognition decisions categorised by the judgment response they later received. For source judgments, Experiments 4.2 and 4.3 did not find a consistent speed advantage for accurate source items but re-analysis of data combined across both experiments demonstrated that recognition was faster for items when the subsequent source judgment was correct. The differing findings across experiments mirrors previous research. Some studies have found a difference between RT for incorrect and correct source items (e.g., Kahn et al., 2004; Lundstrom et al., 2005) and other studies have not (e.g., Wilding & Rugg, 1996). From the combined analysis in the present experiments, it is tentatively concluded that RTs are faster for correct recognition decisions where source judgments are ultimately correct. The retrieval of source information at recognition speeds the recognition decision.

In contrast to the mixed results for source, for categories of subjective experience RT findings were very consistent across experiments. To my knowledge, the current experiments were the first to examine RTs using the four categories of Remember, Know, Familiar and Guess. Here, with the separation of Knowing and Familiarity, recognition decisions for Know items were generally found to be equally fast as for Remember items (Experiments 3.1, 3.3, 4.1, and 4.2; though in the 2AFC task of Experiment 3.3 Remember were found to be significantly faster than Know) and Remember and Know were significantly faster than Familiar. These results therefore can be interpreted as agreeing with the findings of earlier studies that found Remember to be faster than Know, as in those studies Know and Familiar had not been separated (Dewhurst & Conway, 1994; Dewhurst et al., 1998, 2006; Henson et al., 1999; Stretch & Wixted, 1998; Vilberg & Rugg, 2007; Wixted & Stretch, 2004).

However, as discussed earlier (Chapter 3, Section 3.2.3.4), many dual-process models predict that responses based on familiarity should be quicker than those based on recollection as familiarity is considered to be rapid and automatic while recollection is thought to be slower and more effortful (e.g., Jacoby, 1991; Mandler, 1980; Yonelinas & Jacoby, 1994, 1996). This inconsistency between the models' predictions and experimental findings led Yonelinas (2002) to suggest that the finding of Remember responses being faster than Know responses is an artefact of instructions which require participants to respond Know "only if an item is 'familiar and not recollected'. [Thus]… subjects are essentially instructed to wait until both processes are complete before making a know response" (p. 462). If the processes of recollection and familiarity are considered to directly map onto Remember and Familiar subjective experiences then Yonelinas (2002) could be correct – Familiar responses are slower because experimental instructions require the recollection process to be completed first. Furthermore, with the separation of Know and Familiar categories, Yonelinas's suggestion is still able to explain the present findings that Know and Remember RTs were equally fast. If Know is considered to be high confidence without recollection then whether the recollection process results in anything from encoding being recollected (Remember) or nothing being recollected (Know) this should take the same amount of time and recognition RTs for Remembering or Knowing should be equal, as was demonstrated in the current experiments.

However, Dewhurst et al. (2006) provided evidence against the suggestion that Remember responses are faster than Know responses as an artefact of experimental instructions (Yonelinas, 2002). In their experiment, faster recognition decisions for Remember items were demonstrated even when the judgment of subjective experience was decoupled from the recognition decision. Under these conditions, experimental instructions regarding

Remembering and Knowing could not influence the timing of recognition decisions and thus Dewhurst et al. concluded against the demand characteristics account proposed by Yonelinas (2002). Furthermore, Dewhurst et al. suggested that instead of directly reflecting the time-course of the *processes* of recollection and familiarity, Remember and Know responses reflect the time taken to make recognition decisions based on *the information which results from* recollection and familiarity processes. If a test item cues retrieval of contextual information from encoding such as thoughts, images, or associations, then recognition decisions can be made more quickly than if it requires additional processing to evaluate its familiarity relative to other items on the test (Conway & Dewhurst, 1994; Dewhurst et al., 2006; Henson et al., 1999). For the present experiments, this explains the difference in speed of Remember and Familiar decisions but not the similarity in speed of Remember and Know decisions. Furthermore, in Experiment 4.2 for these RT findings both Know and Familiar responses were accompanied by above-chance source accuracy demonstrating that some contextual information from encoding was retrieved for these items, even if this retrieval was not accompanied by a subjective feeling of Remembering.

As Know and Remember were equally fast in terms of RT, and Remember, Know, and Familiar were accompanied by above-chance source accuracy, it is suggested it is the *subjective feeling* concerning recognition that is critically important in influencing the speed with which decisions are made. For Remember and Know responses, whether an item is retrieved with recollection of contextual information from encoding or high confidence without recollection (accompanied by ‘just known’ contextual information or not) results in two different subjective experiences which are experienced equally rapidly and which result in fast recognition decisions. In contrast, when the subjective experience is one of Familiarity, additional processing regarding the source of this feeling of familiarity (and any accompanying contextual information retrieved) results in recognition decisions taking longer. The subjective feeling is related to the information in memory and in what manner context is retrieved, and assessment of all these elements of retrieval influence recognition decision RTs.

For confidence judgments, Experiments 4.1 and 4.3 demonstrated that patterns of recognition RTs categorised by later confidence judgment were similar to those for subjective experience judgments. Recognition RTs for High and Medium confidence items were faster than for Low and None. The general finding that the higher the confidence the faster the response replicates previous research (Henson et al., 2000; Mandler & Boeck, 1974; Ratcliff & Murdock, 1976). However, the similarity between the patterns for confidence and subjective experience RTs suggests that the two types of judgment are similar in terms of RT while they were not found to be similar in terms of distribution of responses or recognition accuracy (see Section 5.2.3). This key thesis theme of the relationship between confidence and subjective experience is discussed further in Section 5.4.

5.2.5. Reaction time to make judgments

This thesis was the first time that reaction times to make post-recognition judgments of confidence, source, and subjective experience have been compared (Experiments 4.1, 4.2, and 4.3). Consequently, this is the first time that source judgments were shown to be faster than confidence judgments, which in turn were faster than subjective experience judgments. This second finding was predicted, the first was not. As both source and subjective experience judgments require assessment of what information can be retrieved regarding context at encoding it was predicted that both source and subjective experience judgments would be slower than judgments of confidence. One suggestion for the observed pattern is that judgments of subjective experience and confidence are both subjective types of assessment whereas source is objective. Thus, confidence and subjective experience judgments may require greater metacognitive processing because of their non-objective nature. Additionally, source judgments may have been so quick because only a one-dimensional source judgment was used and associative encoding had been performed at study. These manipulations led to source judgments being accurate over 80% of the time and may have also influenced RTs. As discussed In Chapter 4 (Section 4.4.4), future experiments could examine whether source judgments are slower when the source judgment is more difficult, or when the delay between encoding and retrieval is longer, as this might increase forgetting of source information. Previous research has shown that coarse source discriminations are faster than finer ones (Lindsay, 2008). For example, Johnson, Kounios, and Reeder (1994) analysed the time-course of item recognition and source memory and found that item recognition became accurate more quickly than source memory. However, multi-dimensional source experiments have not yet investigated the time course of source judgments. If this was explored in future research it might be expected that, in line with the findings of Meiser and Sattler (2007; see Section

5.2.3), RTs might differ for source judgments which require retrieval of specific (e.g., individual voice) and global (e.g., gender of voice) information.

In the present experiments, in addition to recognition judgments being faster when a subsequent judgment of source was correct, RTs to make source judgments were also shown to be faster when source was correct. This finding is novel, though unsurprising. It also adds to the findings of Kahn et al. (2004) and Lundstrom et al. (2005) who used one-step methods where recognition and source were judged at the same time. The current findings demonstrate that when separate source and recognition judgments are made, both these types of judgment are faster when source is accurate. Critically, using the two-step method in the current experiments, the speed advantage for correct source judgments demonstrated a larger and more reliable difference than the speed advantage observed for recognition decisions later shown to have correct source. Moreover, further analysis revealed that RTs for source judgments were faster when accompanied by Remember or High confidence. Thus, the subjective experience or confidence associated with recognition of an item was related to how quickly a source judgment was made to that item.

This novel finding is particularly interesting in terms of the relationship between source and subjective experience as for subjective experience judgments themselves no reliable relationship was demonstrated between category of subjective experience and speed of judgment. The subjective experience associated with recognition of an item did not influence RTs to make subjective experience judgments yet it did influence RTs for source judgments. This is surprising due to the finding that source judgments were made significantly faster than subjective experience judgments. Though source judgments were made more quickly, within that quick response there was significant variation and this was linked to the subjective experience with which the item was recognised. It is suggested that this occurred because a judgment of source is an easier one-dimensional type of judgment compared to a subjective experience judgment, at least in these experiments. Thus, if the critical source information was recollected from encoding and this was experienced subjectively as Remembering then this led to speeding of the source judgment. This interplay between source and subjective experience fits with the suggestion made earlier (Section 5.2.4) that recognition decision RTs themselves are based on the relationship between the subjective experience of recognition and the manner in which contextual information accompanying them is retrieved.

In contrast to the finding that source judgments were speeded by accompanying recollection, Remember responses themselves were not made significantly more quickly than other subjective experience responses. One explanation for this could be that all processing of subjective experience has already taken place at recognition and judgment RT simply reflects time taken to make a response. However, as subjective experience judgments took longer to make than recognition decisions (1794ms compared to 1560ms respectively[36]), this suggests that some additional processing was being undertaken at the time of making the judgment. Participants could have been checking that they did not in fact retrieve any contextual information; or checking the content of the contextual information they had retrieved – was it accurate? Metacognitive processing such as this would result in the patterns observed; as soon as processing ends a response is made, and RT for this response is not dependent on the outcome of the processing. This suggestion fits with the pattern of results obtained for confidence judgments where judgments were found to be made more quickly as confidence increased. For these more inferential, less subjective judgments (see Section 5.4), it is suggested that the type of processing required for these judgments is easier than that for subjective experience judgments. Although the subjective experience with which an item is recognised is suggested to be what gives rise to feelings of confidence (Gardiner, 2002; Tulving, 1985; Section 5.4), once subjective experience has been triggered by presentation of an item at test, the lack of further processing required post-recognition results in quicker judgments when confidence is higher.

To summarise the key take home points from each chapter, the experiments presented in Chapter 2 established that reports of others' subjective experiences are interpreted, in part, by the confidence they convey but that the underlying experiential state is understood as being more important than the confidence for conceptualisations of subjective experience. Chapter 3's experiments demonstrated that recollection can be used to overcome confusion arising from experimental familiarity but that when the relative familiarity of items is manipulated within-subjects subjective experience usage

[36] Average RTs calculated across experiments. RTs for Guess excluded from calculations.

carries-over to those items which are not familiar. In Chapter 4, experiments demonstrated that subjective experience and confidence judgments are not used in the same manner and both have different relationships with the objective measure of source accuracy. Across Chapters 3 and 4, recognition and judgment RTs also provided evidence of differences between source, confidence, and subjective experience. These findings are discussed in relation to the key thesis themes of the separation of Knowing and Familiarity, the differences between confidence and subjective experience, and the influence of objective manipulations on subjective experience in the following sections.

5.3. The separation of Know and Familiar

The key issue within the separation of Know and Familiar categories of subjective experience is how these categories should be defined and understood. What does a *Know* subjective experience entail? What does a *Familiar* subjective experience entail? And in what ways are these categories of subjective experience similar and, more importantly, different? The current thesis was the first to employ the four categories of Remember, Know, Familiar, and Guess in simple item and 2AFC recognition paradigms. Previous research has only employed the separation of Know and Familiar with objectively more meaningful stimuli such as learning of rare word definitions and student course material where the topic of interest was the semanticisation of knowledge over time (e.g., Conway et al., 1997; Dewhurst et al., 2009; Herbert & Burt, 2001, 2003, 2004). In these paradigms, separation of Know and Familiar makes intuitive sense as the Know category is seen as capturing the subjective experience of retrieving information from the semantic memory system once it has been integrated with other knowledge. A key question for the experiments in this thesis was therefore: Would participants use the Know response in the episodic memory tasks? And furthermore, would there be reliable differences between recognition responses assigned to Know and recognition responses assigned to Familiar? Findings from the present experiments that speak to these points are summarised in Table 5.2.

Table 5.2. Main findings for the separated categories of Know and Familiar.

Factor	Know	Familiar
Recognition accuracy	K was recognised just as accurately as R (Exp. 3.1, 3.2[1], 3.3, 4.1, 4.2)	F was recognised less accurately than K (Exp. 3.1, 3.2[2], 3.3[3], 4.1, 4.2,)
Recognition RT	K was recognised just as quickly as R (Exp. 3.1, 3.3[4], 4.1, 4.2)	F was recognised more slowly than K (Exp. 3.1, 3.3, 4.1, 4.2)
Source accuracy	Source judgments for K were less accurate than for R (exp. 4.2)	Source judgments for F were less accurate than for K (Exp. 4.2)
Manipulation of familiarity	Pre-exposure led to reduced K (due to increased R, Exp. 3.1). K was less accurate in Pre-Exposure than in No Pre-Exposure (Exp. 3.3 Yes/No task only). K was not correlated with memory performance (Exp. 3.1, 3.2, 3.3)	F was less accurate in Pre-Exposure than in No Pre-Exposure (Exp. 3.1, 3.2, 3.3[5]). F was negatively correlated with memory performance (Exp. 3.1 – Pre-Exposure, 3.2, 3.3)
Manipulation of confidence	Assignment of K justifications was influenced by associated confidence level (Exp. 2.2)	F justifications were consistently assigned to F (Exp. 2.2)
Confidence	K justifications were given lower confidence ratings than R (Exp. 2.1). K and R items did not differ in terms of confidence (Exp. 4.1)	F justifications were given lower confidence ratings than K (Exp. 2.1). F items were given lower confidence ratings than K (Exp. 4.1)
Sorting task	K justifications were reliably sorted to K	F justifications were not reliably sorted. F were rated as containing more information and details than K

[1] Except in No Pre-Exposure
[2] Except in Pre-Exposure
[3] Except in No Pre-Exposure in 2AFC task
[4] Except in the 2AFC task
[5] Except in the Yes/No task

In terms of similarities, the current experiments suggest that Know and Familiar experiences are similar in how often they are reported by participants. Experiments 4.1 and 4.2 demonstrated very similar *a priori* usage of Know and Familiar responses by participants; approximately 25% of correct recognition responses were assigned each to Know and Familiar. However, this is where the similarity ends. In the recognition memory experiments of this thesis Know recognition was demonstrated to be more accurate (Experiments 3.1, 3.2, 3.3, 4.1, and 4.2), more confident (Experiments 4.1), faster (Experiments 3.1, 3.3, 4.1, and 4.2), and have higher source accuracy (Experiment 4.2) than Familiar recognition. In fact, in the majority of analyses Know recognition was found to be no different to Remember recognition on all of the above factors. The only factor that was shown to reliably differentiate Remember and Know responses was the source accuracy associated with the response (Experiment 4.2). That there were no differences in accuracy or confidence between Remember and Know recognition responses replicates the results of Dewhurst et al. (2009). From the current experiments, the novel findings regarding source accuracy and RT add to Dewhurst et al.'s results regarding how Remember, Know, and Familiar subjective experiences differ.

Reliable differences for Know and Familiar were also observed in Experiment 3.1 where familiarity of items was manipulated through pre-exposure. The between-subjects design of Experiment 3.1 demonstrated an experimental dissociation between Know and Familiar as participants made more Remember responses and fewer Know responses in the Pre-Exposure condition compared to the No Pre-Exposure condition while proportion of responses assigned to Familiar did not differ across conditions. When recollection was required to overcome the familiarity induced by pre-exposure the increase in Remember responses was accompanied by a decrease in Know responses, not a decrease in Familiar responses; indeed, all items would have felt familiar, they had been pre-exposed. The use of familiarity when items had been pre-exposed was primarily what led to the reduction in memory performance in Experiments 3.1 to 3.3. Familiar responses were associated with reduced recognition accuracy and a negative relationship with memory performance when any or all items had been pre-exposed. In contrast, there was no relationship between Know responses and memory performance and the accuracy of Know responses did not differ across conditions.

These results establish that participants are able to use the Know response in episodic memory tasks. A subjective experience of 'just knowing' that an item was presented previously without any associated recollection of context or feelings of familiarity was how a significant proportion of items were recognised in the current experiments. Previous research had only used the separate categories of Know and Familiar with more meaningful stimuli (Conway et al., 1997; Dewhurst et al., 2009; Herbert & Burt, 2001, 2003, 2004). The current experiments demonstrate that the separation of Know and Familiar is also valid for episodic recognition paradigms and use of these separate categories can result in valuable patterns of data which may have been obscured if a conflated Know-Familiar response option had been utilised.

Further differences between Know and Familiar were also demonstrated by Experiments 2.1 to 2.4, which examined lay-person interpretations of memory justification statements. Know justifications received higher confidence ratings (Experiment 2.1), were rated as containing fewer details (Experiment 2.4), were sorted to their category more reliably (Experiment 2.4), and were assigned to the appropriate response less consistently when confidence was manipulated (Experiment 2.2), compared to Familiar justifications. As discussed in Chapter 2 (Sections 2.4.4 and 2.6.1) it is suggested that Know justifications reflected a subjective feeling of high confidence occurring without recollection whereas Familiar justifications reflected feelings of familiarity. The different bases of these subjective experiences (as articulated in the justifications) are suggested to be what led to the differences in findings in the experiments in Chapter 2.

That Know subjective experiences are conceived as reflecting high confidence without recollection and Familiar subjective experiences are conceived as reflecting feelings of familiarity with an item are not themselves surprising given that this was how these types of experience were defined for participants in Experiments 3.1, 3.2, 3.2, 4.1, and 4.2 (See Appendix A). However, the findings regarding recognition accuracy, RT, and confidence obtained for Know and Familiar responses in these experiments provide objective support to these conceptualisations of Know and Familiar. Moreover, the confidence ratings, sorting patterns, and information ratings obtained for Know and Familiar justifications in Experiments 2.1 and 2.4 demonstrated that even when no definitions of subjective experience categories were provided to participants, Know and Familiar responses were able to be differentiated. Critically, the bases for this differentiation also reflected the same differences in confidence and lack of recollection that comprised the experimental definitions. Know justifications were given higher confidence ratings and were considered to contain fewer memory details (indicating an absence of recollection) compared to Familiar justifications.

That Know justifications are considered to reflect high confidence without recollection also explains the findings of Experiment 2.2. When Know justifications that had received differing confidence ratings in Experiment 2.1 were presented for classification, how these were assigned to subjective experience categories was greatly influenced by the associated confidence level. As these justifications were all about confidence and lacked recollective details or mentioned that these were absent, if the confidence reflected in the justifications was low participants often classified the justification as reflecting a subjective experience associated with a lower level of confidence such as Familiar or Guess. In contrast, Familiar justifications of any level of confidence were consistently assigned to Familiar suggesting that the experiential state reflected in these justifications was understood by participants even when the confidence associated with that experiential state differed.

There are two main points to take from these results. First, participants are able to use the Know response in recognition tasks with objectively less-meaningful stimuli such as words and names; and second, reliable differences are observed between items recognised with feelings of Knowing and feelings of Familiarity. As suggested by Gardiner (2001) concerning the traditional Remember-Know paradigm, "Other classifications of subjective conscious experiences of memory are undoubtedly possible and may even prove more useful, in the longer run, than this one." (p. 1360). The separation of Know and Familiar is suggested to be one of these more useful classifications and one that may benefit future research in the areas of aging, neuropsychology, education, and in identifying the neurological bases of memory retrieval. The usefulness of the Know-Familiar split for each of these areas is discussed in Section 5.7.

5.4. The dissociation of confidence and subjective experience

With regard to the relationship between subjective experience and confidence, experiments from this thesis support two conclusions from previous research: Confidence and subjective experience are related; however, they are not one and the same. As already discussed (see Sections 5.2.1 and 5.2.3), Experiments 2.1 and 4.1 demonstrated that levels of confidence decline across Remember, Know, Familiar and Guess experiences; this was found both when it was a first-person assessment of one's own memory under study (Experiment 4.1) and when it was a third-person assessment concerning a justification statement provided by someone else (Experiment 2.1). Critically however, Experiment 4.1 also demonstrated that participants did not use confidence and subjective experience judgments in the same way. Participants were much more lenient with their confidence judgments; over 60% of correctly recognised items were assigned to High confidence while only 40% of items were assigned to Remember. Moreover, participants did not map the two judgment types onto one another – not all Low confidence responses were assigned to Familiar, not all Medium to Know, not all High to Remember; the judgment types and the levels within those judgment types were not interchangeable (Rajaram et al., 2002).

A further critical dissociation of subjective experience and confidence is evident by comparing across experiments in Chapter 4. Experiments 4.2 and 4.3 employed a source manipulation at encoding where participants were instructed to associate words presented on one side of the screen with one person and words on the other side with a different person. In these experiments, confidence and subjective experience judgments demonstrated different relationships with source accuracy. Subjective experience had a relationship with source when source was both accurate and inaccurate. When source was accurate, more responses were assigned to Remember, followed by Know, followed by Familiar; in contrast, when source was inaccurate more responses were assigned to Familiar, then Know, then Remember. For confidence judgments, a relationship only existed between confidence and source when source was accurate; more responses were assigned to High, then Medium, then Low confidence. For inaccurate source judgments, equal proportions were assigned to High, Medium, and Low confidence. As discussed in Section 5.2.3, this was taken as indicating that subjective experience judgments are more sensitive and more reflective of memorial processes than judgments of confidence. It is hypothesised that both source and subjective experience judgments require assessment of whether contextual details have been retrieved, whereas this assessment is not required when a confidence judgment is made.

Further dissociation of confidence and subjective experience is demonstrated by the fact that the source encoding manipulation employed in Experiments 4.2 and 4.3 led to increased recognition performance and higher confidence in these experiments compared to Experiment 4.1, but patterns of subjective experience did not change. This finding runs contrary to previous manipulations, which have demonstrated that when more elaborative or conceptual processing is performed at encoding more Remember responses are made. Such manipulations include deep versus shallow processing, generation or vocalization versus reading of items (Gardiner, 1988; Gregg & Gardiner, 1991), orthographically distinctive versus orthographically common words (Rajaram, 1998), repetition of items at study (Dewhurst & Anderson, 1999), and intentional versus incidental learning (Macken & Hampson, 1993). It is unclear why the source encoding manipulation employed in

Experiments 4.2 and 4.3 led to increases in recognition performance and confidence but no increases in Remember responses. In interpreting this result, it is important to remember that this difference was observed across experiments, not within. But the fact that confidence increases but Remembering does not suggests that the two processes are dissociable in some tasks and in response to some variables.

To fully explore how subjective experience and confidence judgments respond to source encoding instruction manipulations, source encoding would need to be performed at study and both subjective experience and confidence judgments made at test. Using this design, it might be predicted that confidence judgments would not be so high when made in conjunction with subjective experience judgments. Though this might seem counterintuitive, this prediction is suggested as experiments in this thesis have demonstrated that subjective experience judgments are more sensitive to manipulations of source than confidence judgments are (see Section 5.2.3); and experiments in this thesis and previous research have demonstrated that judgments of confidence are more lenient than judgments of subjective experience (e.g., Gardiner & Java, 1990; Rajaram, 1993; Rajaram et al., 2002; and Tunney & Fernie, 2007). As confidence is suggested to be derived from subjective experience (Gardiner, 2001; Tulving, 1985), it is suggested that a source encoding manipulation combined with subjective experience judgments at retrieval might lead confidence judgments to be adjusted to be in line with the accompanying subjective experience; i.e., rates of High confidence would be lowered to the levels shown in Experiment 4.1.

That subjective experience gives rise to confidence is suggested to be what led to the results of Experiment 2.3. In that experiment, participants were provided with justification statements accompanied by confidence values that were either appropriate or inappropriate to that justification. Tulving asserts that "…the adaptive value of episodic memory and autonoetic consciousness lies in the heightened subjective certainty with which organisms endowed with such memory and consciousness believe, and are willing to act upon, information retrieved from memory" (1985; p. 10). Results of Experiment 2.3 demonstrated that participants did not take into consideration confidence values when assigning justification statements to a category of subjective experience. Although this does not directly test the hypothesis that subjective experience gives rise to confidence, it does demonstrate that experiential reports are considered of more importance than confidence in determining another's recollective status.

On the whole, this thesis interprets confidence as an inferential assessment of subjective experiences, whereas subjective experience judgments themselves are more reflective of the underlying memory processes and experiential states. Furthermore, the current experiments have demonstrated confidence and subjective experience to dissociate in many different ways. These findings are therefore interpreted as supporting dual-process as opposed to single-process accounts of recognition. In agreement with Dewhurst et al. (2006), subjective experience responses in the current experiments are not considered to be direct reflections of the underlying processes of recollection and familiarity, rather they are considered to be assessments of the information provided by these processes *plus* the experiential state which results. Although some of the present findings may be able to be interpreted using unidimensional SDT models, that Remember and Know responses were differentiated only by source accuracy while Know and Familiar responses were dissociated by many factors adds to evidence from other behavioural and brain imaging studies that subjective experience judgments differ in ways other than just confidence or familiarity (Dewhurst et al., 2006; and see Gardiner, 2008, and Gardiner & Richardson-Klavenh, 2000, for reviews).

The results of this thesis may best be interpreted in terms of recent conceptualisations of recollection as a continuous process (e.g., Mickes et al., 2009, 2010; Onyper et al., 2010; Parks & Yonelinas, 2007; Rotello et al., 2004; Slotnick, 2010; Slotnick & Dodson, 2005; Wixted, 2007, 2010; Wixted & Mickes, 2010; Wixted & Stretch, 2004). With the separation of Know and Familiar responses, the finding that Know and Remember responses differ only in terms of source accuracy while Remember, Know, and Familiar responses were all accompanied by retrieval of varying degrees of source is taken to support the idea that there are varying degrees of recollection. Critically, it is not suggested that retrieval of source for Know and Familiar items was accompanied by recollective experience, instead the suggestion is that retrieval of source was experienced either as semantic knowledge (Know) or with feelings of familiarity (Familiar). In addition, the level of accuracy of this retrieval distinguished the different subjective experiences.

While conceptualising recollection as a continuous process may benefit interpretation of experimental findings in terms of the best-fit model, these models are primarily concerned with modelling the recognition decision process and not with theoretically understanding the subjective experience that accompanies recognition (Dewhurst et al., 2006). As Gardiner (2008) highlights:

> "There has recently been a spate of formal quantitative models. But the increasing technical sophistication and complexity of some of these models and the rather general ability of most of them to provide a reasonably good fit to the data make it increasingly difficult to see how to distinguish between them empirically… Confronted by a plethora of alternative versions of such models, it is hard (despite the claims sometimes made for this approach) to see any great advantage of quantitative modelling over a less mathematically, more conceptually driven approach" (pp. 301-302).

The current thesis concentrated on examining how people make and understand judgments of subjective experience and interpreted findings in terms of psychological explanations of human memory. By using this approach, the current experiments demonstrated confidence and subjective experience to be differentiated in a variety of ways, on a variety of different tasks; such differences could not be observed using quantitative modelling approaches.

5.5. How do objective manipulations affect subjective experience?

The effects of objective manipulations on subjective experience have long been studied. As discussed in the General Introduction (Section 1.6.1), some experimental manipulations have been found to increase Remember responses while Know responses are unchanged (e.g., Dewhurst & Anderson, 1999; Gardiner, 1988; Gregg & Gardiner, 1991; Jones & Roediger, 1995; Macken & Hampson, 1993; Rajaram, 1998); others have demonstrated situations where Know responses increase while Remember responses are unaffected (Dewhurst & Hitch, 1997; Mäntylä & Raudsepp, 1996; Rajaram, 1993; Rajaram & Geraci, 2000). Other manipulations have been found to result in opposite effects on Remembering and Knowing (Gardiner & Java, 1990; Mäntylä, 1997; Parkin & Russo, 1993). This last pattern is what was obtained when familiarity was manipulated in Experiment 3.1. When pre-exposure increased the familiarity of items this led to increased Remember responses and a corresponding decrease in Know responses. Critically, this change in patterns of subjective experiences was also accompanied by a reduction in memory performance. When items were pre-exposed, Remembering increased but recognition accuracy decreased.

The different patterns of Remember-Know responses obtained using different manipulations have been taken as support for the 'functional independence' of Remembering and Knowing and support for dual-process accounts of memory (Gardiner, 2008; Gardiner & Conway, 1999). The current findings further support these conclusions. Although single-process protagonists such as Dunn (2004, 2008) have demonstrated that SDT models can account for some previous Remember-Know dissociations it is suggested that the results of Experiment 3.1 present a challenge to SDT models. In this experiment, the 'trace strength' of both targets and lures was increased in the pre-exposure condition but the strength of targets should still have been higher than that of lures as targets were studied subsequent to pre-exposure. UVSD models may be able to explain the memory performance finding as pre-exposure of lures would have increased the variance of these items compared to non-pre-exposed lures and thus the difference between the variances of lure and target items would be smaller in the pre-exposure condition. However, that Remember responses are increased in pre-exposure while recognition performance is decreased is unable to be explained by SDT models. As greater recollection is assumed to go hand in hand with better memory performance, Conway et al. (2001) proposed that SDT interpretations are refuted if the proportion of Remember responses differs between two conditions while the overall hit rate remains the same. Dunn (2004) argued against this and stated that just because hit rates are equal across conditions does not imply that Remember rates should be equal across conditions. However, the current experiment poses a different challenge to SDT models; here memory performance was lower in the condition where more responses were assigned to Remember. From a first-person dual-process standpoint this finding is interpreted as reflecting the use of recollection to overcome the familiarity induced by pre-exposure. How single-process models would explain this finding is open for debate.

5.6. METHODOLOGY IN SUBJECTIVE EXPERIENCE RESEARCH

Experiments in the current thesis have provided evidence which warrants discussion regarding three topics of methodology: between- versus within-subjects designs, one-step versus two-step paradigms, and measuring reaction time to make judgments.

Regarding between- and within-subjects designs, the experiments of Chapter 3 demonstrated that while the effect of pre-exposure on memory performance was evident across all experimental designs, its effects on subjective experience were only evident when pre-exposure was manipulated between-subjects. Previous research has also demonstrated that in some cases experimental design can be critical for demonstrating effects on subjective experience (e.g., Conway et al., 2001; Dewhurst & Parry, 2000). However, to my knowledge, the current experiments are the first to demonstrate a situation where the effect of a manipulation on recognition performance is consistent across experimental designs but subjective experience responses are affected. This underscores the benefit of investigating the relationship between subjective experience and performance using multiple experimental designs.

In relation to one-step versus two-step subjective experience paradigms, while the current thesis did not compare these directly, separation of recognition decisions from post-recognition judgments enabled some key analysis to be performed which could not have been conducted if one-step procedures had been used. Firstly, separating recognition decisions from judgments in Chapter 4 permitted examination of the influence that making two types of judgment for an item might have on responses in a clearer fashion than would have been able if one of the judgments had been combined with recognition. Indeed, results of these experiments demonstrated no effects of ‘contamination’ of one judgment by the other in the patterns of responding, which challenges the marginal evidence of contamination demonstrated by Martin (2007) who had used one-step procedures to try to investigate this hypothesis.

Secondly, the two-step procedures in these experiments enabled measurement of how long it took participants to make subjective judgments of Remember, Know, Familiar, and Guess. Separation of recognition and judgment here meant that the time course of both these decision processes could be analysed separately and that RTs for source, confidence, and subjective experience judgments could be examined free from contamination by recognition decision time. Critically, this novel methodology also enabled the influence of one judgment on the other judgment's RT to be examined and this demonstrated that the speed at which source judgments are made was influenced by the accompanying subjective experience. For this comparison, if either the source or subjective experience judgment had been conflated with recognition then this RT finding may have been obscured by recognition decision time. Fundamentally, the separation of recognition decision processes and judgments means that findings regarding subjective experience, source, and confidence, and the time courses of these types of assessment, are easier to interpret.

5.7. Applications and future research

The findings presented in this thesis have implications for the advancement of theory and methodology in recognition memory research (see Sections 5.3 to 5.5). More importantly, they may also be able to benefit research in aging, neuropsychology, and education, as well as research exploring the neurological bases of memory retrieval.

To take each of these areas in turn, in research into memory in aging, utilising separate Know and Familiar response options may help elucidate whether the increase in Know responses for older adults observed using the traditional Remember-Know paradigm actually results from increases in subjective experiences of Knowing or Familiarity. As discussed in Chapter 1 (Section 1.6.2), some experiments have demonstrated that the reduction in Remembering in aging is accompanied by an increase in Knowing (e.g., Parkin & Walter, 1992; Perfect et al., 1995) while others have found that older adults make fewer Remember responses than younger adults but that level of Know responses do not differ between age groups (e.g., Mäntylä, 1993). Employing the separate categories of Know and Familiar in experimental manipulations may enable clarification of findings from experimental manipulations that have produced different patterns of responding in aging using the traditional Remember-Know paradigm. The separation of Know and Familiar would allow examination of whether it is patterns of Know or Familiar responses that are altered when fewer Remember responses are made. As semantic memory is generally found to remain intact in normal aging while episodic memory declines (Craik, 1999), it might be that reductions in Remember responses are accompanied by increases in Know responses for older adults but that level of Familiar responses remains the same across age groups. As it is recollective experience which declines in aging it follows that increases would be demonstrated in Know responses rather than Familiar responses as Know reflect high confidence *in the absence of recollection* while Familiar responses reflect feelings of familiarity for an item in comparison to other items on the test.

The separation of Know and Familiar would also permit this type of exploration in conditions such as Alzheimer's disease, amnesia, schizophrenia, and autism spectrum disorders. In these populations, while levels of Remember responses are reduced, levels of Know responses are generally found to be unaffected or are affected to a much lesser extent than Remembering (Bowler et al., 2000; Dalla Barba, 1993, 1997; Huron et al., 1995; Knowlton & Squire, 1995; Schacter et al., 1997; Tanweer et al., 2009). The separation of Know and Familiar may aid clarification of which non-recollective states of awareness are spared in different populations and whether it is Knowing or Familiarity (or both) that are spared in particular experimental tasks. Furthermore, separation of these categories may be able to dissociate what types of memory are able to be improved in training paradigms such as that of Jennings and Jacoby (2003) which has examined training recollection in older adults.

As discussed in Chapter 1 (Section 1.6.3), the Remember-Know dissociations observed in neuropsychological populations have been linked to the underlying changes in neurological functioning and/or anatomy for these groups. The separation of Know and Familiar may aid clarification of which non-recollective states of awareness are associated with particular patterns of neuronal activation in both neuropsychologically impaired and normal functioning. Previous evidence regarding the neural substrates of recollection and familiarity is somewhat mixed. For example, when damage to the hippocampus is somewhat selective, some comparisons of recall and recognition performance have shown recollection to be disrupted while familiarity is relatively spared (e.g., Aggleton & Shaw, 1996; Aggleton et al., 2000, 2005; Baddeley et al., 2001; Bowles et al., 2010; Hanley et al., 1994; Holdstock, et al., 2000; Jäger et al., 2009; Mayes et al., 2001, 2002; Turriziani et al., 2008; Vargha-Khadem et al., 1997; Yonelinas et al., 2002); however, others have reported impairments in both recollection and familiarity (Cipolotti, et al., 2006; Gold et al., 2006; Manns et al., 2003; Wais et al., 2006). Additionally, left anterior prefrontal regions have been demonstrated to be more active for Remembered items compared to Known items in some studies but not in others (cf. Henson et al., 1999, Eldridge et al., 2000).

Recent research by Montaldi et al. (2006) demonstrated different patterns of neural activation for different levels of familiarity across a variety of brain areas, but hippocampal activity was not modulated by familiarity. In this study, participants responded using the four categories of Recollect, Very familiar, Moderately familiar, and Slightly familiar. Although these different response options do not correspond to the conceptualisations of Know, Familiar, and Guess in the current thesis, Montaldi et al.'s results do suggest that the neural substrates of recollection and familiarity are not yet clearly delineated. Further evidence of this comes from Diana et al. (2010), who demonstrated that retrieval of different types of source information were related to different patterns of MTL activation. They suggested that when assessing the roles of neural regions in episodic memory, instead of considering activation to be solely determined by memory strength or the processes of recollection and familiarity, "the types of memorial information involved at all stages of processing should be considered" (pp. 1816-1817). This relates to the current findings regarding the relationships between source, subjective experience, and confidence and the suggestion that recollection is a continuous process. These examples demonstrate that the neural substrates of subjective experience are not yet fully understood. As stated by Gardiner (2008, p. 302), "Gaining a better understanding of remembering and knowing theoretically will depend on further evidence that links these states of awareness not only with behaviour but also with the brain". It is suggested that utilising the separate categories of Know and Familiar in both behavioural and neuroimaging studies will aid this theoretical understanding.

The inspiration for the current body of work was the demonstration of an R-to-K shift in learning of academic material by Conway et al. (1997). Conway and colleagues were the first to separate the subjective experiences of Knowing and Familiarity and through this separation were able to show that material shifted from episodic to semantic memory over time. Subsequent experiments by Herbert and Burt (2001, 2003, 2004) replicated these findings using different academic learning materials and Dewhurst et al. (2009) demonstrated that even less objectively meaningful materials demonstrated semanticisation across learning. The current experiments were the first to employ the separate categories of Know and Familiar in standard item and associative recognition memory paradigms and findings demonstrate that this separation is valid for these paradigms. Thus, the subjective experience responses of Remember, Know, Familiar, and Guess appear to be reflective of how items are retrieved from memory across many types of learning task. This extended Remember-Know paradigm is applicable to episodic immediately tested single-item memory experiments, multi-trial learning of semantic information, and longitudinal studies of learning in academic settings. Future research into how subjective experiences differ across different types of material and different methods of learning would be of value. Indeed, the initial study by Conway et al. showed that proportion of Remember and Know responses differed across academic courses; in lecture courses Remember responses were made more often whereas on a research methods course the dominant response was Know. Conway et al. suggested that the development of conceptualised knowledge structures in the research methods course was encouraged by the presentation of material in multiple contexts and the more interactive learning environment. Additionally, Herbert and Burt (2003) demonstrated that students were more likely to shift from Remembering to Knowing if the learning material was reviewed regularly and in different formats. The findings from these studies suggest that systematic examination of the R-to-K shift across different materials and different methods of learning could benefit teaching practises.

5.8. Conclusions

This thesis examined how people make and understand judgments of subjective experience and it was the first to employ the four categories of Remember, Know, Familiar, and Guess in a standard episodic memory task. The manipulations and experimental procedures utilised in the current experiments enabled critical differences between these categories to be demonstrated. Know and Familiar judgments were shown to dissociate on recognition accuracy, source accuracy, confidence, and response time. In contrast, Remember and Know judgments were only able to be differentiated by source accuracy. Results from the present experiments are able to inform psychological conceptualisations of how people understand and judge subjective experience and thus further advancement of theoretical models of memory processes. Looking at the wider picture, better psychological understanding of subjective experience and its relationship to recognition processes is of value to future research in a wide number of domains.

REFERENCES

Åberg, C., & Nilsson, L. G. (2001). Facilitation of source discrimination in the novelty effect. *Scandinavian Journal of Psychology, 42*(4), 349-357.

Åberg, C. S., & Nilsson, L. G. (2003). A strict response criterion yields a mirror effect in the novelty paradigm. *Scandinavian Journal of Psychology, 44*(5), 425-432.

Aggleton, J. P., McMackin, D., Carpenter, K., Hornak, J., Kapur, N., Halpin, S., et al. (2000). Differential cognitive effects of colloid cysts in the third ventricle that spare or compromise the fornix. *Brain, 123*(1), 800-815.

Aggleton, J. P., Seralynne, D. V., Denby, C., Dix, S., Mayes, A. R., Roberts, N., et al. (2005). Sparing of the familiarity component of recognition memory in a patient with hippocampal pathology. *Neuropsychologia, 43*, 1810-1823.

Aggleton, J. P., & Shaw, C. (1996). Amnesia and recognition memory: A re-analysis of psychometric data. *Neuropsychologia, 34*(1), 51-62.

Arndt, J., & Reder, L. A. (2002). Word frequency and receiver operating characteristic curves in recognition memory: Evidence for a dual-process interpretation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(5), 830-842.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89-195). Oxford, England: Academic Press.

Baddeley, A., Vargha-Khadem, F., & Mishkin, M. (2001). Preserved recognition in a case of developmental amnesia: Implications for the acquisition of semantic memory? *Journal of Cognitive Neuroscience, 13*(3), 357-369.

Balota, D. A., & Neely, J. H. (1980). Test-expectancy and word-frequency effects in recall and recognition. *Journal of Experimental Psychology: Human Learning and Memory, 6*(5), 576-587.

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*(8), 934-946.

Bastin, C., & Van der Linden, M. (2003). The contribution of recollection and familiarity to recognition memory: A study of the effects of test format and aging. *Neuropsychology, 17*(1), 14-24.

Blaxton, T. A., & Theodore, W. H. (1997). The role of the temporal lobes in recognizing visuospatial materials: Remembering versus knowing. *Brain and Cognition, 35*, 5-25.

Bodner, G. E., & Lindsay, D. S. (2003). Remembering and knowing in context. *Journal of Memory and Language, 48*(3), 563-580.

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*(7), 718-729.

Bowler, D. M., Gardiner, J. M., & Grice, S. J. (2000). Episodic memory and remembering in adults with Asperger Syndrome. *Journal of Autism and Developmental Disorders, 30*(4), 295-304.

Bowles, B., Crupi, C., Pigott, S., Parrent, A., Wiebe, S., Janzen, L., et al. (2010). Selective recollection impairments following unilateral stereotactic amygdalo-hippocampotomy for treatment of temporal-lobe epilepsy. *Neuropsychologia, 48*(9), 2640-2647.

Brewer, W. F., & Sampaio, C. (2006). Processes leading to confidence and accuracy in sentence recognition: A metamemory approach. *Memory, 14*(5), 540-552.

Bruno, D., & Rutherford, A. (2010). How many response options? A study of remember-know testing procedures. *Acta Psychologica, 134*(2), 125-129.

Cary, M., & Reder, L. M. (2003). A dual-process account of the list-length and strength-based mirror effects in recognition. *Journal of Memory and Language, 49*(2), 231-248.

Chalmers, D. J. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies, 2*(3), 200-219.

Chalmers, K. A., & Humphreys, M. S. (1998). Role of generalized and episode specific memories in the word frequency effect in recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(3), 610-632.

Cipolotti, L., Bird, C., Good, T., Macmanus, D., Rudge, P., & Shallice, T. (2006). Recollection and familiarity in dense hippocampal amnesia: A case study. *Neuropsychologia, 44*, 489-506.

Clark, S. E. (1992). Word-frequency effects in associative and item recognition. *Memory & Cognition, 20*(3), 231-243.

Clark, S. E., & Burchett, R. E. R. (1994). Word-frequency and list composition effects in associative recognition and recall. *Memory & Cognition, 22*(1), 55-62.

Clark, S. E., & Gronlund, S. D. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin & Review, 3*(1), 37-60.

Cohen, G. (1981). Inferential reasoning in old-age. *Cognition, 9*(1), 59-72.

Conway, M. A. (2005). Memory and the self. *Journal of Memory and Language, 53*(4), 594-628.

Conway, M. A., Collins, A. F., Gathercole, S. E., & Anderson, S. J. (1996). Recollections of true and false autobiographical memories. *Journal of Experimental Psychology: General, 125*, 69-95.

Conway, M. A., & Dewhurst, S. A. (1995a). The self and recollective experience. *Applied Cognitive Psychology, 9*(1), 1-19.

Conway, M. A., & Dewhurst, S. A. (1995b). Remembering, familiarity, and source monitoring. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 48*(1), 125-140.

Conway, M. A., Dewhurst, S. A., Pearson, N., & Sapute, A. (2001). The self and recollection reconsidered: how a 'failure to replicate' failed and why trace strength accounts of recollection are untenable. *Applied Cognitive Psychology, 15*(6), 673-686.

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*(4), 393-413.

Conway, M. A., & Williams, H. L. (2008). The nature of autobiographical memory. In J. H. Byrne (Ed.), *Learning and memory: A comprehensive reference* (pp. 893-909). Oxford: Elsevier Ltd.

Cook, G. I., Marsh, R. L., & Hicks, J. L. (2005). Revisiting the role of recollection in item versus forced-choice recognition memory. *Psychonomic Bulletin & Review, 12*(4), 720-725.

Craik, F. I. M. (1999). Age-related changes in human memory. In D. Park & N. Schwarz (Eds.), *Cognitive aging: A primer* (pp. 75–92). Hove, UK: Psychology Press.

Craik, F. I. M., Morris, L. W., Morris, R. G., & Loewen, E. R. (1990). Relations between source amnesia and frontal lobe functioning in older adults. *Psychology and Aging, 5*(1), 148-151.

Crawley, S., & French, C. (2005). Field and observer viewpoint in remember-know memories of personal childhood events. *Memory, 13*(7), 673 - 681.

Curran, T. (2004). Effects of attention and confidence on the hypothesized ERP correlates of recollection and familiarity. *Neuropsychologia, 42*, 1088-1106.

Curran, T., & Doyle, J. (2011). Picture superiority doubly dissociates the ERP correlates of recollection and familiarity. *Journal of Cognitive Neuroscience, 23*(5), 1247-1262.

Curran, T., Tepe, K. L., & Piatt, C. (2006). ERP explorations of dual processes in recognition memory. In H. D. Zimmer, A. Mecklinger & U. Lindenberger (Eds.), *Binding in human memory: A neurocognitive approach* (pp. 467-492). Oxford: Oxford University Press.

Dalla Barba, G. (1993). Confabulation: Knowledge and recollective experience. *Cognitive Neuropsychology, 10*(1), 1-20.

Dalla Barba, G. (1997). Recognition memory and recollective experience in Alzheimer's disease. *Memory, 5*, 657-672.

Daniels, K. A., Toth, J. P., & Hertzog, C. (2009). Aging and recollection in the accuracy of judgments of learning. *Psychology and Aging, 24*(2), 494-500.

Davachi, L., Mitchell, J., & Wagner, A. (2003). Multiple routes to memory: Distinct medial temporal lobe processes build item and source memories. *Proceedings of the National Academy of Sciences, 100*, 2157-2162.

DeLosh, E. L., & McDaniel, M. A. (1996). The role of order information in free recall: Application to the word-frequency effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(5), 1136-1146.

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

Dewhurst, S. A., & Conway, M. A. (1994). Pictures, images, and recollective experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(5), 1088-1098.

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*(6), 849-858.

Dewhurst, S. A., & Hitch, G. J. (1997). Illusions of familiarity caused by cohort activation. *Psychonomic Bulletin & Review, 4*, 566-571.

Dewhurst, S. A., & Hitch, G. J. (1999). Cognitive effort and recollective experience in recognition memory. *Memory, 7*(2), 129-146.

Dewhurst, S. A., Hitch, G. J., & Barry, C. (1998). Separate effects of word frequency and age of acquisition in recognition and recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(2), 284-298.

Dewhurst, S. A., Holmes, S. J., Brandt, K. R., & Dean, G. M. (2006). Measuring the speed of the conscious components of recognition memory: Remembering is faster than knowing. *Consciousness and Cognition, 15*(1), 147-162.

Dewhurst, S. A., & Parry, L. A. (2000). Emotionality, distinctiveness, and recollective experience. *European Journal of Cognitive Psychology, 12*(4), 541-551.

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

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2008). The effects of unitization on familiarity-based source memory: Testing a behavioral prediction derived from neuroimaging data. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 730-740.

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2010). Medial temporal lobe activity during source retrieval reflects information type, not memory strength. *Journal of Cognitive Neuroscience, 22*(8), 1808-1818.

Dobbins, I. G., Kroll, N. E. A., & Liu, Q. (1998). Confidence-accuracy inversions in scene recognition: A remember-know analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(5), 1306-1315.

Dobbins, I. G., Kroll, N. E. A., Yonelinas, A. P., & Liu, Q. (1998). Distinctiveness in recognition and free recall: The role of recollection in the rejection of the familiar. *Journal of Memory and Language, 38*, 381-400.

Dodson, C. S., & Johnson, M. K. (1993). Rate of false source attributions depends on how questions are asked. *American Journal of Psychology, 106*(4), 541-557.

Donaldson, W. (1996). The role of decision processes in remembering and knowing. *Memory & Cognition, 24*(4), 523-533.

Donaldson, W., MacKenzie, T. M., & Underhill, C. F. (1996). A comparison of recollective memory and source monitoring. *Psychonomic Bulletin & Review, 3*(4), 486-490.

Dosher, B. A. (1984). Discriminating preexperimental (semantic) from learned (episodic) associations: A speed-accuracy study. *Cognitive Psychology, 16*(4), 519-555.

Duchek, J. M., & Neely, J. H. (1989). A dissociative word-frequency X levels-of-processing interaction in episodic recognition and lexical decision tasks. *Memory & Cognition, 17*(2), 148-162.

Dudukovic, N. M., & Knowlton, B. J. (2006). Remember-Know judgments and retrieval of contextual details. *Acta Psychologica, 122*(2), 160-173.

Düzel, E., Yonelinas, A. P., Mangun, G. R., Heinze, H. J., & Tulving, E. (1997). Event-related brain potential correlates of two states of conscious awareness in memory. *Proceedings of the National Academy of Sciences of the United States of America, 94*(11), 5973.

Dunn, J. C. (2004). Remember-know: A matter of confidence. *Psychological Review, 111*(2), 524-542.

Dunn, J. C. (2008). The dimensionality of the remember-know task: A state-trace analysis. *Psychological Review, 115*(2), 426-446.

Ebbinghaus, H. E. (1885). *Memory: A contribution to experimental psychology*. Republished 1964, New York: Dover.

Eichenbaum, H., Yonelinas, A. R., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience, 30*, 123-152.

Eldridge, L. L., Knowlton, B. J., Furmanski, C. S., Bookheimer, S. Y., & Engel, S. A. (2000). Remembering episodes: a selective role for the hippocampus during retrieval. *Nature Neuroscience, 3*(11), 1149-1152.

Eldridge, L. L., Sarfatti, S., & Knowlton, B. J. (2002). The effect of testing procedure on remember-know judgments. *Psychonomic Bulletin & Review, 9*(1), 139-145.

Fabiani, M., & Friedman, D. (1996). Dissociations between memory for temporal order and recognition memory in aging. *Neuropsychologia, 35*(2), 129-141.

Fife-Shaw, C. (2005). Calculator Utility. Retrieved 1st February, 2009, from http://www.surrey.ac.uk/psychology/current/statistics/

Fisher, I. (1921). The best form of index number. *Quarterly Publications of the American Statistical Association, 17*, 533-551.

Friedman, D., & Johnson, R. (2000). Event-related potential (ERP) studies of memory encoding and retrieval: A selective review. *Microscopy Research and Technique, 51*(1), 6-28.

Friedman, D., & Trott, C. (2000). An event-related potential study of encoding in young and older adults. *Neuropsychologia, 38*(5), 542-557.

Gardiner, J. M. (1988). Functional aspects of recollective experience. *Memory & Cognition, 16*(4), 309-313.

Gardiner, J. M. (2000). On the objectivity of subjective experiences of autonoetic and noetic consciousness. In E. Tulving (Ed.), *Memory, consciousness, and the brain: The Tallinn conference* (pp. 159-172). Philadelphia: Psychology Press.

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

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

Gardiner, J. M., & Conway, M. A. (1999). Levels of awareness and varieties of experience. In B. H. Challis & B. M. Velichkovsy (Eds.), *Stratification of consciousness and cognition* (pp. 237-254). Amsterdam/Philadelphia: Benjamin.

Gardiner, J. M., & Gregg, V. H. (1997). Recognition memory with little or no remembering: Implications for a detection model. *Psychonomic Bulletin & Review, 4*(4), 474-479.

Gardiner, J. M., & Java, R. I. (1990). Recollective experience in word and nonword recognition. *Memory & Cognition, 18*(1), 23-30.

Gardiner, J. M., Java, R. I., & Richardson-Klavehn, A. (1996). How level of processing really influences awareness in recognition memory. *Canadian Journal of Experimental Psychology, 50*(1), 114-122.

Gardiner, J. M., Kaminska, Z., Dixon, M., & Java, R. I. (1996). Repetition of previously novel melodies sometimes increases both remember and know responses in recognition memory. *Psychonomic Bulletin & Review, 3*(3), 366-371.

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

Gardiner, J. M., Ramponi, C., & Richardson-Klavehn, A. (1999). Response deadline and subjective awareness in recognition memory. *Consciousness and Cognition, 8*(4), 484-496.

Gardiner, J. M., Ramponi, C., & Richardson-Klavehn, A. (2002). Recognition memory and decision processes: A meta-analysis of remember, know, and guess responses. *Memory, 10*(2), 83-98.

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

Gardiner, J. M., Richardson-Klavehn, A., & Ramponi, C. (1997). On reporting recollective experiences and ''direct access to memory systems''. *Psychological Science, 8*(5), 391-394.

Geraci, L., McCabe, D. P., & Guillory, J. J. (2009). On interpreting the relationship between remember-know judgments and confidence: The role of instructions. *Consciousness and Cognition, 18*(3), 701-709.

Geraci, L., & McCabe, D. R. (2006). Examining the basis for illusory recollection: The role of remember/know instructions. *Psychonomic Bulletin & Review, 13*(3), 466-473.

Gilhooly, K. J., & Logie, R. H. (1980). Age of acquisition, imagery, concreteness, familiarity and ambiguity measures for 1944 words. *Behavior Research Methods & Instrumentation, 12*(4), 395-427.

Gillund, G., & Shiffrin, R. M. (1984). A Retrieval Model for Both Recognition and Recall. *Psychological Review, 91*(1), 1-67.

Glanzer, M., & Adams, J. K. (1990). The mirror effect in recognition memory: Data and theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(1), 5.

Glanzer, M., & Bowles, N. (1976). Analysis of the word-frequency effect in recognition memory. *Journal of Experimental Psychology: Human Learning and Memory, 2*(1), 21-31.

Glanzer, M., Kim, K., Hilford, A., & Adams, J. K. (1999). Slope of the receiver-operating characteristic in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(2), 500-513.

Gold, J. J., Smith, C. N., Bayley, P. J., Shrager, Y., Brewer, J. B., Stark, C. E. L., et al. (2006). Item memory, source memory, and the medial temporal lobe: Concordant findings from fMRI and memory-impaired patients. *Proceedings of the National Academy of Sciences, 103*(24), 9351-9356.

Goodman, L. A., & Kruskal, W. H. (1954). Measures of association for cross classifications. *Journal of the American Statistical Association, 49*(268), 732-764.

Greene, R. L. (1999). The role of familiarity in recognition. *Psychonomic Bulletin & Review, 6*(2), 309-312.

Gregg, V. (1976). Word frequency, recognition and recall. In J. Brown (Ed.), *Recall and recognition* (pp. 183-216). London: Wiley.

Gregg, V. H., & Gardiner, J. M. (1991). Components of conscious awareness in a long-term modality effect. *British Journal of Psychology, 82*, 153-162.

Gregg, V. H., Montgomery, D. C., & Castano, D. (1980). Recall of common and uncommon words from pure and mixed lists. *Journal of Verbal Learning and Verbal Behavior, 19*(2), 240-245.

Gronlund, S. D., & Ratcliff, R. (1989). Time course of item and associative information: Implications for global memory models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(5), 846-858.

Gruppuso, V., Lindsay, D. S., & Kelley, C. M. (1997). The process-dissociation procedure and similarity: Defining and estimating recollection and familiarity in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(2), 259-278.

Gruppuso, V., Lindsay, D. S., & Masson, M. E. J. (2007). I'd know that face anywhere! *Psychonomic Bulletin & Review, 14*(6), 1085-1089.

Guttentag, R. E., & Carroll, D. (1997). Recollection-based recognition: Word frequency effects. *Journal of Memory and Language, 37*(4), 502-516.

Hacker, M. J., & Ratcliff, R. (1979). Revised table of d' for m-alternative forced choice. *Perception & Psychophysics, 26*(2), 168-170.

Hanley, J. R., Davies, A. D. M., Downes, J. J., & Mayes, A. R. (1994). Impaired recall of verbal material following rupture and repair of an anterior communicating artery aneurysm. *Cognitive Neuropsychology, 11*(5), 543-578.

Haug, H., & Eggers, R. (1991). Morphometry of the human cortex cerebri and corpus striatum during aging. *Neurobiology of Aging, 12*(4), 336-338.

Henke, K., Weber, B., Kneifel, S., Wieser, H. G., & Buck, A. (1999). Human hippocampus associates information in memory. *Proceedings of the National Academy of Sciences of the United States of America, 96*(10), 5884-5889.

Henson, R. N. A., Rugg, M. D., Shallice, T., & Dolan, R. J. (2000). Confidence in recognition memory for words: Dissociating right prefrontal roles in episodic retrieval. *Journal of Cognitive Neuroscience, 12*(6), 913-923.

Henson, R. N. A., Rugg, M. D., Shallice, T., Josephs, O., & Dolan, R. J. (1999). Recollection and familiarity in recognition memory: An event-related functional magnetic resonance imaging study. *The Journal of Neuroscience, 19*(10), 3962-3972.

Herbert, D. M. B., & Burt, J. S. (2001). Memory awareness and schematization: Learning in the university context. *Applied Cognitive Psychology, 15*(6), 617-637.

Herbert, D. M. B., & Burt, J. S. (2003). The effects of different review opportunities on schematisation of knowledge. *Learning and Instruction, 13*(1), 73-92.

Herbert, D. M. B., & Burt, J. S. (2004). What do students remember? Episodic memory and the development of schematization. *Applied Cognitive Psychology, 18*(1), 77-88.

Hicks, J. L., & Marsh, R. L. (1999). Remember-know judgments can depend on how memory is tested. *Psychonomic Bulletin & Review, 6*(1), 117-122.

Hicks, J. L., Marsh, R. L., & Ritschel, L. (2002). The role of recollection and partial information in source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(3), 503-508.

Hintzman, D. L. (2001). Judgments of frequency and recency: How they relate to reports of subjective awareness. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(6), 1347-1358.

Hintzman, D. L., & Curran, T. (1994). Retrieval dynamics of recognition and frequency judgments: Evidence for separate processes of familiarity and recall. *Journal of Memory and Language, 33*(1), 1-18.

Hirshman, E., & Lanning, K. (1999). Is there a special association between self judgements and conscious recollection? *Applied Cognitive Psychology, 13*(1), 29-42.

Hirshman, E., & Master, S. (1997). Modeling the conscious correlates of recognition memory: Reflections on the remember-know paradigm. *Memory & Cognition, 25*(3), 345-351.

Hirst, W., Johnson, M. K., Phelps, E. A., & Volpe, B. T. (1988). More on recognition and recall in amnesics. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(4), 758-762.

Hockley, W. E. (1984). Retrieval of item frequency information in a continuous memory task. *Memory & Cognition, 12*(3), 229-242.

Hockley, W. E., & Consoli, A. (1999). Familiarity and recollection in item and associative recognition. *Memory & Cognition, 27*(4), 657-664.

Holdstock, J. S., Mayes, A. R., Cezayirli, E., Isaac, C. L., Aggleton, J. P., & Roberts, N. (2000). A comparison of egocentric and allocentric spatial memory in a patient with selective hippocampal damage. *Neuropsychologia, 38*(4), 410-425.

Holmes, J. B., Waters, H. S., & Rajaram, S. (1998). The phenomenology of false memories: Episodic content and confidence. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 1026-1040.

Humphreys, M. S. (1978). Item and relational information: Case for context independent retrieval. *Journal of Verbal Learning and Verbal Behavior, 17*(2), 175-187.

Humphreys, M. S., Dennis, S., Maguire, A. M., Reynolds, K., Bolland, S. W., & Hughes, J. D. (2003). What you get out of memory depends on the question you ask. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(5), 797-812.

Huron, C., Danion, J. M., Giacomoni, F., Grange, D., Robert, P., & Rizzo, L. (1995). Impairment of recognition memory with, but not without, conscious recollection in Schizophrenia. *American Journal of Psychiatry, 152*(12), 1737-1742.

Inoue, C., & Bellezza, F. S. (1998). The detection model of recognition using know and remember judgments. *Memory & Cognition, 26*(2), 299-308.

Isaac, C. L., & Mayes, A. R. (1999). Rate of forgetting in amnesia: I. Recall and recognition of prose. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(4), 942-962.

Jacoby, L. L. (1983). Remembering the data: Analyzing interactive processes in reading. *Journal of Verbal Learning and Verbal Behavior, 22*(5), 485-508.

Jacoby, L. L. (1984). Incidental versus intentional retrieval: Remembering and awareness as separate issues. In L. R. Squire & N. Butters (Eds.), *Neuropsychology of memory* (pp. 145-156). New York: Guilford Press.

Jacoby, L. L. (1991). A process dissociation framework - separating automatic from intentional uses of memory. *Journal of Memory and Language, 30*(5), 513-541.

Jacoby, L. L. (1999). Ironic effects of repetition: Measuring age-related differences in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(1), 3-22.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 110*(3), 306-340.

Jacoby, L. L., & Kelley, C. M. (1992). Unconscious influences of memory: Dissociations and automaticity. In A. D. Milner & M. D. Rugg (Eds.), *The neuropsychology of consciousness* (pp. 201-233). San Diego, CA: Academic Press.

Jacoby, L. L., Kelley, C. M., & Dywan, J. (1989). Memory attributions. In H. L. Roediger III & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 391-422). Hillsdale, NJ: Erlbaum.

Jacoby, L. L., & Witherspoon, D. (1982). Remembering without awareness. *Canadian Journal of Psychology, 36*(2), 300-324.

Jacoby, L. L., Yonelinas, A. P., & Jennings, J. M. (1997). The relation between conscious and unconscious (automatic) influences: A declaration of independence. In J. D. Cohen &

J. W. Schooler (Eds.), *Scientific approaches to consciousness* (pp. 13-47). Hillsdale, NJ: Erlbaum.

Jäger, T., Szabo, K., Griebe, M., Bäzner, H., Möller, J., & Hennerici, M. G. (2009). Selective disruption of hippocampus-mediated recognition memory processes after episodes of transient global amnesia. *Neuropsychologia, 47*(1), 70-76.

James, W. (1890). *Principles of psychology* (Vol. 1). New York: Holt.

Java, R. I. (1996). Effects of age on state of awareness following implicit and explicit word-association tasks. *Psychology and Aging, 11*(1), 108-111.

Jennings, J. M., & Jacoby, L. L. (2003). Improving memory in older adults: Training recollection. *Neuropsychological Rehabilitation, 13*(4), 417-440.

Johansson, M., Mecklinger, A., & Treese, A. C. (2004). Recognition memory for emotional and neutral faces: An event-related potential study. *Journal of Cognitive Neuroscience, 16*(10), 1840-1853.

Johnson, M. K., & Kim, J. K. (1985). Recognition of pictures by alcoholic Korsakoff patients. *Bulletin of the Psychonomic Society, 23*(6), 456-458.

Johnson, M. K., Kounios, J., & Reeder, J. A. (1994). Time-course studies of reality monitoring and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(6), 1409.

Jones, T. C., & Roediger, H. L., III. (1995). The experiential basis of serial position effects. *European Journal of Cognitive Psychology, 7*(1), 65-80.

Joordens, S., & Hockley, W. E. (2000). Recollection and familiarity through the looking glass: When old does not mirror new. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(6), 1534-1555.

Kahn, I., Davachi, L., & Wagner, A. D. (2004). Functional-neuroanatomic correlates of recollection: Implications for models of recognition memory. *The Journal of Neuroscience, 24*(17), 4172-4180.

Kelley, C. M., & Jacoby, L. L. (1990). The construction of subjective experience: Memory attributions. *Mind & Language, 5*(1), 49-68.

Kelley, C. M., & Jacoby, L. L. (1998). Subjective reports and process dissociation: Fluency, knowing, and feeling. *Acta Psychologica, 98*(2-3), 127-140.

Kensinger, E. A., & Schacter, D. L. (2006). Amygdala activity is associated with the successful encoding of item, but not source, information for positive and negative stimuli. *The Journal of Neuroscience, 26*(9), 2564-2570.

Khoe, W., Kroll, N. E. A., Yonelinas, A. P., Dobbins, I. G., & Knight, R. T. (2000). The contribution of recollection and familiarity to yes-no and forced-choice recognition tests in healthy subjects and amnesics. *Neuropsychologia, 38*(10), 1333-1341.

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

Kilb, A., & Naveh-Benjamin, M. (2007). Paying attention to binding: Further studies assessing the role of reduced attentional resources in the associative deficit of older adults. *Memory & Cognition, 35*(5), 1162-1174.

Kinoshita, S. (1995). The word-frequency effect in recognition memory versus repetition priming. *Memory & Cognition, 23*(5), 569-580.

Kinsbourne, M., & George, J. (1974). The mechanism of the word-frequency effect on recognition memory. *Journal of Verbal Learning and Verbal Behavior, 13*(1), 63-69.

Knowlton, B. J., & Squire, L. R. (1995). Remembering and knowing: two different expressions of declarative memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(3), 699-710.

Komatsu, S., Graf, P., & Uttl, B. (1995). Process dissociation procedure: Core assumptions fail, sometimes. *European Journal of Cognitive Psychology, 7*(1), 19-40.

Kormi-Nouri, R., Nilsson, L. G., & Ohta, N. (2005). The novelty effect: Support for the Novelty-Encoding Hypothesis. *Scandinavian Journal of Psychology, 46*(2), 133-143.

Kroll, N. E. A., Yonelinas, A. P., Dobbins, I. G., & Frederick, C. M. (2002). Separating sensitivity from response bias: Implications of comparisons of yes-no and forced-choice tests for models and measures of recognition memory. *Journal of Experimental Psychology: General, 131*(2), 241-254.

Kronlund, A., & Whittlesea, B. W. A. (2005). Seeing double: Levels of processing can cause false memory. *Canadian Journal of Experimental Psychology, 59*(1), 11-16.

Kronlund, A., & Whittlesea, B. W. A. (2006). Remembering after a perception of discrepancy: Out with the old, in with the two. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(5), 1174-1184.

Lee, K., & Wixted, J. T. (2004). Signal Detection Theory. Retrieved 1st February, 2009, from http://psy2.ucsd.edu/~kang/sdt/sdt.htm

Lindsay, D. S. (2008). Source monitoring. In J. H. Byrne (Ed.), *Learning and memory: A comprehensive reference* (pp. 325-348). Oxford: Elsevier Ltd.

Lundstrom, B. N., Ingvar, M., & Petersson, K. M. (2005). The role of precuneus and left inferior frontal cortex during source memory episodic retrieval. *Neuroimage, 27*(4), 824-834.

Macken, W. J., & Hampson, P. (1993). Integration, elaboration and recollective experience. *Irish Journal of Psychology, 14*, 279-285.

MacLeod, C. M., & Kampe, K. E. (1996). Word frequency effects on recall, recognition, and word fragment completion tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(1), 132-142.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Macmillan, N. A., & Rotello, C. M. (2006). Deciding about decision models of remember and know judgments: A reply to Murdock (2006). *Psychological Review, 113*, 657-665.

Macmillan, N. A., Rotello, C. M., & Verde, M. (2005). On the importance of models in interpreting remember-know experiments: Comments on Gardiner et al.'s (2002) meta-analysis. *Memory, 13*(6), 607-621.

Maddox, W. T., & Estes, W. K. (1997). Direct and indirect stimulus-frequency effects in recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(3), 539-559.

Mandler, G. (1979). Organization and repetition: Organizational principles with special reference to rote learning. In L. G. Nilsson (Ed.), *Perspectives on memory research* (pp. 293-327). Hillsdale, NJ: Erlbaum.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review, 87*(3), 252-271.

Mandler, G. (1991). Your face looks familiar but I can't remember your name: A review of dual process theory. In W. E. Hockley & S. Lewandowsky (Eds.), *Relating theory and data: Essays on human memory in honor of Bennet B. Murdock* (pp. 207-225). Hillsdale, NJ: Erlbaum.

Mandler, G. (2008). Familiarity breeds attempts: A critical review of dual process theories of recognition. *Perspectives on Psychological Science, 3*(5), 390-399.

Mandler, G., & Boeck, W. J. (1974). Retrieval processes in recognition. *Memory & Cognition, 2*(4), 613-615.

Mandler, G., Pearlstone, Z., & Koopmans, H. S. (1969). Effects of organization and semantic similarity on recall and recognition. *Journal of Verbal Learning and Verbal Behavior, 8*(3), 410-423.

Manns, J. R., Hopkins, R. O., Reed, J. M., Kitchener, E. G., & Squire, L. R. (2003). Recognition memory and the human hippocampus. *Neuron, 37*(1), 171-180.

Mäntylä, T. (1993). Knowing but not remembering: Adult age differences in recollective experience. *Memory & Cognition, 21*(3), 379-388.

Mäntylä, T. (1997). Recollections of faces: Remembering differences and knowing similarities. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(5), 1203-1216.

Mäntylä, T., & Raudsepp, J. (1996). Recollective experience following suppression of focal attention. *European Journal of Cognitive Psychology, 8*(2), 195-204.

Marsh, R. L., & Hicks, J. L. (1998). Test formats change source-monitoring decision processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(5), 1137-1151.

Martin, B. A. (2007). *Source memory, subjective awareness, and the word frequency mirror effect.* Unpublished Masters, Louisiana State University and Agricultural and Mechanical College.

Mayes, A. R., Holdstock, J. S., Isaac, C. L., Hunkin, N. M., & Roberts, N. (2002). Relative sparing of item recognition memory in a patient with adult onset damage limited to the hippocampus. *Hippocampus, 12*(3), 325-340.

Mayes, A. R., Isaac, C. L., Holdstock, J. S., Hunkin, N. M., Montaldi, D., Downes, J. J., et al. (2001). Memory for single items, word pairs, and temporal order of different kinds in a patient with selective hippocampal lesions. *Cognitive Neuropsychology, 18*(2), 97-123.

McCabe, D. P., & Balota, D. A. (2007). Context effects on remembering and knowing: The expectancy heuristic. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(3), 536-549.

McCabe, D. P., & Geraci, L. D. (2009). The influence of instructions and terminology on the accuracy of remember-know judgments. *Consciousness and Cognition, 18*(2), 401-413.

McElree, B., Dolan, P. O., & Jacoby, L. L. (1999). Isolating the contributions of familiarity and source information to item recognition: A time course analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(3), 563.

Mecklinger, A. (2000). Interfacing mind and brain: A neurocognitive model of recognition memory. *Psychophysiology, 37*(05), 565-582.

Mecklinger, A. (2006). Electrophysiological measures of familiarity memory. *Journal of Clinical EEG & Neuroscience, 37*(4), 292-299.

Meiser, T., & Bröder, A. (2002). Memory for multidimensional source information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(1), 116-137.

Meiser, T., & Sattler, C. (2007). Boundaries of the relation between conscious recollection and source memory for perceptual details. *Consciousness and Cognition, 16*(1), 189-210.

Mickes, L., Johnson, E. M., & Wixted, J. T. (2010). Continuous recollection versus unitized familiarity in associative recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(4), 843-863.

Mickes, L., Wais, P. E., & Wixted, J. T. (2009). Recollection is a continuous process. *Psychological Science, 20*(4), 509-515.

Mirandola, C., Del Prete, F., Ghetti, S., & Cornoldi, C. (2011). Recollection but not familiarity differentiates memory for text in students with and without learning difficulties. *Learning and Individual Differences, 21*, 206-209.

Montaldi, D., Spencer, T. J., Roberts, N., & Mayes, A. R. (2006). The neural system that mediates familiarity memory. *Hippocampus, 16*(5), 504-520.

MRC Psycholinguistic Database: Machine Usable Dictionary. from http://www.psych.rl.ac.uk/MRC_Psych_Db.html

Mulligan, N. W., & Hirshman, E. (1997). Measuring the bases of recognition memory: An investigation of the process-dissociation framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(2), 280-304.

Murdock, B. B. (2006). Decision-making models of remember-know judgments: Comment on Rotello, Macmillan, and Reeder (2004). *Psychological Review, 113*, 648-656.

National Health Service Register. (2007). Most Common Surnames in England, Wales and the Isle of Man. Retrieved 1st October, 2008, from www.surnames.behindthename.com

Naveh-Benjamin, M. (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(5), 1170-1187.

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin, 95*(1), 109-133.

Nolde, S. F., Johnson, M. K., & D'Esposito, M. (1998). Left prefrontal activation during episodic remembering: An event-related fMRI study. *Neuroreport, 9*(15), 3509-3514.

Norman, K. A., & Schacter, D. L. (1997). False recognition in younger and older adults: Exploring the characteristics of illusory memories. *Memory & Cognition, 25*(6), 838-848.

Nyberg, L., Cabeza, R., & Tulving, E. (1996). PET studies of encoding and retrieval: The HERA model. *Psychonomic Bulletin & Review, 3*(2), 135-148.

Office of National Statistics. (2007). Baby Names Statistics Girls/Boys. Retrieved 1st October, 2008, from www.statistics.gov.uk

Old, S. R., & Naveh-Benjamin, M. (2008). Differential effects of age on item and associative measures of memory: A meta-analysis. *Psychology and Aging, 23*(1), 104-118.

Onyper, S. V., Zhang, Y. X., & Howard, M. W. (2010). Some-or-none recollection: Evidence from item and source memory. *Journal of Experimental Psychology: General, 139*(2), 341-364.

Palmer, M. A., Brewer, N., McKinnon, A. C., & Weber, N. (2010). Phenomenological reports diagnose accuracy of eyewitness identification decisions. *Acta Psychologica, 133*(2), 137-145.

Parkin, A. J., & Russo, R. (1993). On the origin of functional differences in recollective experience. *Memory, 1*(3), 231-237.

Parkin, A. J., & Walter, B. M. (1992). Recollective experience, normal aging, and frontal dysfunction. *Psychology and Aging, 7*(2), 290-298.

Parks, C. M., & Yonelinas, A. P. (2007). Moving beyond pure signal-detection models: Comment on Wixted (2007). *Psychological Review, 114*(1), 188-201.

Parks, C. M., & Yonelinas, A. P. (2008). Theories of recognition memory. In J. H. Byrne (Ed.), *Learning and memory: A comprehensive reference* (pp. 389-416). Oxford: Elsevier Ltd.

Perfect, T. J., & Dasgupta, Z. R. R. (1997). What underlies the deficit in reported recollective experience in old age? *Memory & Cognition, 25*(6), 849-858.

Perfect, T. J., Mayes, A. R., Downes, J. J., & Van Eijk, R. (1996). Does context discriminate recollection from familiarity in recognition memory? *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 49*(3), 797-813.

Perfect, T. J., Williams, R. B., & Anderton-Brown, C. (1995). Age differences in reported recollective experience are due to encoding effects, not response bias. *Memory, 3*(2), 169-186.

Rajaram, S. (1993). Remembering and knowing: Two means of access to the personal past. *Memory & Cognition, 21*(1), 89-102.

Rajaram, S. (1998). The effects of conceptual salience and perceptual distinctiveness on conscious recollection. *Psychonomic Bulletin & Review, 5*(1), 71-78.

Rajaram, S., & Geraci, L. (2000). Conceptual fluency selectively influences knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(4), 1070-1074.

Rajaram, S., Hamilton, M., & Bolton, A. (2002). Distinguishing states of awareness from confidence during retrieval: Evidence from amnesia. *Cognitive Affective & Behavioral Neuroscience, 2*(3), 227-235.

Ramponi, C., Barnard, P., & Nimmo-Smith, I. (2004). Recollection deficits in dysphoric mood: An effect of schematic models and executive mode? *Memory, 12*(5), 655-670.

Ranganath, C., Johnson, M. K., & D'Esposito, M. (2000). Left anterior prefrontal activation increases with demands to recall specific perceptual information. *The Journal of Neuroscience, 20*(22), RC108.

Ranganath, C., Yonelinas, A. P., Cohen, M. X., Dy, C. J., Tom, S. M., & D'Esposito, M. (2004). Dissociable correlates of recollection and familiarity within the medial temporal lobes. *Neuropsychologia, 42*(1), 2-13.

Ratcliff, R., & Murdock, B. B. (1976). Retrieval processes in recognition memory. *Psychological Review, 83*(3), 190-214.

Raye, C. L., Johnson, M. K., Mitchell, K. J., Nolde, S. F., & D'Esposito, M. (2000). fMRI investigations of left and right PFC contributions to episodic remembering. *Psychobiology, 28*(2), 197-206.

Reder, L. M., Nhouyvanisvong, A., Schunn, C. D., Ayers, M. S., Angstadt, P., & Hiraki, K. (2000). A mechanistic account of the mirror effect for word frequency: A computational model of remember-know judgments in a continuous recognition paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(2), 294-320.

Richardson-Klavehn, A., Gardiner, J. M., & Java, R. I. (1996). Memory: Task dissociations, process dissociations, and dissociations of consciousness. In G. F. Underwood (Ed.), *Implicit Cognition* (pp. 85-155). Oxford: Oxford University Press.

Rotello, C. M., & Heit, E. (2000). Associative recognition: A case of recall-to-reject processing. *Memory & Cognition, 28*(6), 907-922.

Rotello, C. M., & Macmillan, N. A. (2006). Remember-know models as decision strategies in two experimental paradigms. *Journal of Memory and Language, 55*(4), 479-494.

Rotello, C. M., Macmillan, N. A., & Reeder, J. A. (2004). Sum-difference theory of remembering and knowing: A two-dimensional signal-detection model. *Psychological Review, 111*(3), 588-616.

Rotello, C. M., Macmillan, N. A., Reeder, J. A., & Wong, M. (2005). The remember response: Subject to bias, graded, and not a process-pure indicator of recollection. *Psychonomic Bulletin & Review, 12*(5), 865-873.

Rotello, C. M., & Zeng, M. (2008). Analysis of RT distributions in the remember-know paradigm. *Psychonomic Bulletin & Review, 15*(4), 825-832.

Rugg, M. D., & Allan, K. (2000). Event-related potential studies of memory. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 521-537).

Rugg, M. D., Cox, C. J. C., Doyle, M. C., & Wells, T. (1995). Event-related potentials and the recollection of low and high-frequency words. *Neuropsychologia, 33*(4), 471-484.

Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences, 11*(6), 251-257.

Rugg, M. D., Schloerscheidt, A. M., & Mark, R. E. (1998). An electro-physiological comparison of two indices of recollection. *Journal of Memory and Language, 39*, 47-69.

Salthouse, T. A. (1991). Mediation of adult age differences in cognition by reductions in working memory and speed of processing. *Psychological Science, 2*(3), 179-183.

Schacter, D. L., Verfaellie, M., & Anes, M. D. (1997). Illusory memories in amnesic patients: Conceptual and perceptual false recognition. *Neuropsychology, 11*(3), 331-342.

Schacter, D. L., Verfaellie, M., & Pradere, D. (1996). The neuropsychology of memory illusions: False recall and recognition in amnesic patients. *Journal of Memory and Language, 35*(2), 319-334.

Senkfor, A. J., & Van Petten, C. (1998). Who said what? An event-related potential investigation of source and item memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(4), 1005-1025.

Skinner, E. I., & Fernandes, M. A. (2007). Neural correlates of recollection and familiarity: A review of neuroimaging and patient data. *Neuropsychologia, 45*(10), 2163-2179.

Slotnick, S. D. (2010). “Remember” source memory ROCs indicate recollection is a continuous process. *Memory, 18*(1), 27-39.

Slotnick, S. D., & Dodson, C. S. (2005). Support for a continuous (single-process) model of recognition memory and source memory. *Memory & Cognition, 33*(1), 151-170.

Smith, M. E. (1993). Neurophysiological manifestations of recollective experience during recognition memory judgments. *Journal of Cognitive Neuroscience, 5*(1), 1-13.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117*(1), 34-50.

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*(4), 769-784.

Souchay, C., Williams, H. L., Moulin, C. J. A., Isingrini, M., & Conway, M. A. (2006, July). *Monitoring judgments and recollection in older adults.* Paper presented at the International Conference on Memory (ICOM-4), Sydney, Australia.

Spencer, W. D., & Raz, N. (1994). Memory for facts, source, and context: Can frontal lobe dysfunction explain age-related differences? *Psychology and Aging, 9*(1), 149-159.

Squire, L. R., Wixted, J. T., & Clark, R. E. (2007). Recognition memory and the medial temporal lobe: A new perspective. *Nature Reviews Neuroscience, 8*(11), 872-883.

Starns, J. J., & Hicks, J. L. (2005). Source dimensions are retrieved independently in multidimensional monitoring tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(6), 1213-1220.

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin, 87*(2), 245-251.

Strack, F., & Förster, J. (1995). Reporting recollective experiences: Direct access to memory systems? *Psychological Science, 6*(6), 352-358.

Strange, B. A., Otten, L. J., Josephs, O., Rugg, M. D., & Dolan, R. J. (2002). Dissociable human perirhinal, hippocampal, and parahippocampal roles during verbal encoding. *The Journal of Neuroscience, 22*(2), 523-528.

Tanweer, T., Rathbone, C. J., & Souchay, C. (2009). Autobiographical memory, autonoetic consciousness, and identity in Asperger syndrome. *Neuropsychologia, 48*(4), 900-908.

Tendolkar, I., Arnold, J., Petersson, K. M., Weis, S., Brockhaus-Dumke, A., Van Eijndhoven, P., et al. (2008). Contributions of the medial temporal lobe to declarative memory retrieval: Manipulating the amount of contextual retrieval. *Learning & Memory, 15*(9), 611-617.

Toglia, M. P., & Battig, W. R. (1978). *Handbook of semantic word norms*. New York: Erlbaum.

Trott, C. T., Friedman, D., Ritter, W., Fabiani, M., & Snodgrass, J. G. (1999). Episodic priming and memory for temporal source: Event-related potentials reveal age-related differences in prefrontal functioning. *Psychology and Aging, 14*(3), 390-413.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381-402). New York: Academic Press.

Tulving, E. (1982). Synergistic ecphory in recall and recognition. *Canadian Journal of Psychology, 36*(2), 130-147.

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

Tulving, E. (1995). Organization of memory: Quo vadis? In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 839-847). Cambridge, MA: MIT Press.

Tulving, E., & Kroll, N. (1995). Novelty assessment in the brain and long-term memory encoding. *Psychonomic Bulletin & Review, 2*(3), 387-390.

Tulving, E., & Markowitsch, H. J. (1998). Episodic and declarative memory: role of the hippocampus. *Hippocampus, 8*(3), 198-204.

Tulving, E., Markowitsch, H. J., Craik, F. I. M., Habib, R., & Houle, S. (1996). Novelty and familiarity activations in PET studies of memory encoding and retrieval. *Cerebral Cortex, 6*(1), 71-79.

Tulving, E., Markowitsch, H. J., Kapur, S., Habib, R., & Houle, S. (1994). Novelty encoding networks in the human brain: Positron emission tomography data. *Neuroreport, 5*(18), 2525-2528.

Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science, 247*(4940), 301-306.

Tunney, R. J., & Fernie, G. (2007). Repetition priming affects guessing not familiarity. *Behavioral and Brain Functions, 3*(1), 40.

Turriziani, P., Serra, L., Fadda, L., Caltagirone, C., & Carlesimo, G. A. (2008). Recollection and familiarity in hippocampal amnesia. *Hippocampus, 18*(5), 469-480.

Vargha-Khadem, F., Gadian, D. G., Watkins, K. E., Connelly, A., Van Paesschen, W., & Mishkin, M. (1997). Differential effects of early hippocampal pathology on episodic and semantic memory. *Science, 277*(5324), 376-380.

Verfaellie, M., & Treadwell, J. R. (1993). Status of recognition memory in amnesia. *Neuropsychology, 7*(1), 5-13.

Vilberg, K. L., & Rugg, M. D. (2007). Dissociation of the neural correlates of recognition memory according to familiarity, recollection, and amount of recollected information. *Neuropsychologia, 45*(10), 2216-2225.

Volpe, B. T., Holtzman, J. D., & Hirst, W. (1986). Further characterization of patients with amnesia after cardiac arrest: Preserved recognition memory. *Neurology, 36*(3), 408-411.

Wagner, A. D., Gabrieli, J. D. E., & Verfaellie, M. (1997). Dissociations between familiarity processes in explicit recognition and implicit perceptual memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(2), 305-323.

Wagner, A. D., Koutstaal, W., & Schacter, D. L. (1999). When encoding yields remembering: Insights from event-related neuroimaging. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences, 354*(1387), 1307-1324.

Wais, P. E. (2008). fMRI signals associated with memory strength in the medial temporal lobes: A meta-analysis. *Neuropsychologia, 46*, 3185-3196.

Wais, P. E., Wixted, J. T., Hopkins, R. O., & Squire, L. R. (2006). The hippocampus supports both the recollection and the familiarity components of recognition memory. *Neuron, 49*(3), 459-466.

Weis, S., Specht, K., Klaver, P., Tendolkar, I., Willmes, K., Ruhlmann, J., et al. (2004). Process dissociation between contextual retrieval and item recognition. *Neuroreport, 15*(18), 2729-2733.

Wheeler, M. A., Stuss, D. T., & Tulving, E. (1997). Toward a theory of episodic memory: The frontal lobes and autonoetic consciousness. *Psychological Bulletin, 121*(3), 331-354.

Whittlesea, B. W. A. (2002). Two routes' to remembering (and another to remembering not). *Journal of Experimental Psychology: General, 131*(3), 325-348.

Whittlesea, B. W. A. (2004). The perception of integrality: Remembering through the validation of expectation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(4), 891-908.

Whittlesea, B. W. A., Jacoby, L. L., & Girard, K. (1990). Illusions of immediate memory: Evidence of an attributional basis for feelings of familiarity and perceptual quality. *Journal of Memory and Language, 29*(6), 716-732.

Whittlesea, B. W. A., & Williams, L. D. (2000). The source of feelings of familiarity: The discrepancy-attribution hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(3), 547-565.

Whittlesea, B. W. A., & Williams, L. D. (2001a). The discrepancy-attribution hypothesis: I. The heuristic basis of feelings of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(1), 3-13.

Whittlesea, B. W. A., & Williams, L. D. (2001b). The discrepancy-attribution hypothesis: II. Expectation, uncertainty, surprise, and feelings of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(1), 14-33.

Wilding, E. L., & Rugg, M. D. (1996). An event-related potential study of recognition memory with and without retrieval of source. *Brain, 119*, 889-905.

Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review, 114*(1), 152-176.

Wixted, J. T. (2010, July). *Measuring recollection and familiarity in the medial temporal lobe.* Paper presented at the MRC Recognition Memory Symposium, University of Bristol, Bristol, England.

Wixted, J. T., & Mickes, L. (2010). A continuous dual-process model of Remember/Know judgments. *Psychological Review, 117*(4), 1025-1054.

Wixted, J. T., & Squire, L. R. (2004). Recall and recognition are equally impaired in patients with selective hippocampal damage. *Cognitive, Affective, & Behavioral Neuroscience, 4*(1), 58-66.

Wixted, J. T., & Stretch, V. (2004). In defense of the signal detection interpretation of remember/know judgments. *Psychonomic Bulletin & Review, 11*(4), 616-641.

Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(6), 1341-1354.

Yonelinas, A. P. (1997). Recognition memory ROCs for item and associative information: The contribution of recollection and familiarity. *Memory & Cognition, 25*(6), 747-763.

Yonelinas, A. P. (1999). The contribution of recollection and familiarity to recognition and source-memory judgments: A formal dual-process model and an analysis of receiver operating characteristics. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(6), 1415-1434.

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

Yonelinas, A. P. (2001b). Components of episodic memory: the contribution of recollection and familiarity. *Philosophical Transactions of the Royal Society B-Biological Sciences, 356*(1413), 1363-1374.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46*(3), 441-517.

Yonelinas, A. P., Aly, M., Wang, W. C., & Koen, J. D. (2010). Recollection and familiarity: Examining controversial assumptions and new directions. *Hippocampus, 20*(11), 1178-1194.

Yonelinas, A. P., Dobbins, I., Szymanski, M. D., Dhaliwal, H. S., & King, L. (1996). Signal-detection, threshold, and dual-process models of recognition memory: ROCs and conscious recollection. *Consciousness and Cognition, 5*(4), 418-441.

Yonelinas, A. P., Hopfinger, J. B., Buonocore, M. H., Kroll, N. E. A., & Baynes, K. (2001). Hippocampal, parahippocampal and occipital-temporal contributions to associative and item recognition memory: an fMRI study. *Neuroreport, 12*(2), 359-363.

Yonelinas, A. P., & Jacoby, L. L. (1994). Dissociations of processes in recognition memory: Effects of interference and of response speed. *Canadian Journal of Experimental Psychology, 48*(4), 516-535.

Yonelinas, A. P., & Jacoby, L. L. (1996). Noncriterial recollection: Familiarity as automatic, irrelevant recollection. *Consciousness and Cognition, 5*(1-2), 131-141.

Yonelinas, A. P., Kroll, N. E. A., Dobbins, I., Lazzara, M., & Knight, R. T. (1998). Recollection and familiarity deficits in amnesia: Convergence of remember-know, process dissociation, and receiver operating characteristic data. *Neuropsychology, 12*(3), 323-339.

Yonelinas, A. P., Kroll, N. E. A., Dobbins, I. G., & Soltani, M. (1999). Recognition memory for faces: When familiarity supports associative recognition judgments. *Psychonomic Bulletin & Review, 6*(4), 654-661.

Yonelinas, A. P., Kroll, N. E. A., Quamme, J. R., Lazzara, M. M., Sauvé, M. J., Widaman, K. F., et al. (2002). Effects of extensive temporal lobe damage or mild hypoxia on recollection and familiarity. *Nature Neuroscience, 5*(11), 1236-1241.

Yonelinas, A. P., Otten, L. J., Shaw, K. N., & Rugg, M. D. (2005). Separating the brain regions involved in recollection and familiarity in recognition memory. *The Journal of Neuroscience, 25*(11), 3002-3008.

Yonelinas, A. P., & Parks, C. M. (2007). Receiver operating characteristics (ROCs) in recognition memory: A review. *Psychological Bulletin, 133*(5), 800-832.

Appendix A: Remember, Know, Familiar, and Guess Definitions

Across the experiments presented in this thesis, the definitions of Remember, Know, Familiar and Guess given to participants were kept as consistent as possible. Sometimes wording was changed slightly, particularly when the recognition test was 2AFC as opposed to Yes/No. Definitions from each experiment are shown below.

A.1. Definitions of Remember, Know, Familiar, and Guess from Experiments 2.2 and 2.3.

In these questionnaires the task was for respondents to judge which subjective experience category a previous experimental participant's justification statement should be classed as.

Remember: For this item they had an experience of Remembering the word, this could have included seeing the word in their mind's eye, remembering what they thought or pictured when they saw the word on the original list, and/or having a sense of themselves in the past. *For example, if you see someone on the street you may think 'who is that? Oh yes, I remember, I was in the chemist shop, it's the person I saw in the queue at the chemist, I remember thinking what a funny hat they had on…'*

Know: For this item they simply Know the word without any of the other feelings associated with vividly remembering that they had seen the word before. *For example, if you see someone on the street you may think 'who is that? Oh yes, it's my friend George, I know him really well…'*

Familiar: For this word the person had a feeling of Familiarity with the word and because of that they think that the word was on the previous list. *For example, if you see someone on the street you may think 'who is that? They look very familiar… I don't know where I know them from but they are definitely familiar…'*

Guess: For this word the person had no feeling of familiarity or any other memories associated with the word and simply Guessed that the word was on the previous list.

A.2. Definitions of Remember, Know, Familiar, and Guess from Experiments 3.1 and 3.2.

These experiments involved a 2AFC task where participants had to select which of two forenames had previously been studied with a particular surname. The definitions were altered to reflect this.

R = For this item you have an experience of **Remembering** the name, this could have included seeing the name pair in your mind's eye, remembering what you thought or pictured when you saw the name pair on the original list, and/or having a sense of yourself in the past. *For example, if you see someone on the street you may think 'who is that? Oh yes, I remember, I was in the chemist shop, it's the person I saw in the queue at the chemist, I remember thinking what a funny hat they had on…'*

K = For this item you simply **Know** the first name that was paired with this surname without any of the other feelings associated with vividly remembering that you have seen the item before. *For example, if you see someone on the street you may think 'who is that? Oh yes, it's my friend Rob, I know him really well…'*

F = For this item you have a feeling of **Familiarity** with the first name and because of that you think that that first name had been paired with the surname. *For example, if you see someone on the street you may think 'who is that? They look very familiar… I don't know where I know them from but they are definitely familiar…'*

G = For this item you had no feeling of familiarity or any other memories associated with the name pair and simply **Guessed** that that first name had been paired with that surname.

A.3. Definitions of Remember, Know, Familiar, and Guess from Experiments 3.3, 4.1, and 4.2.

R = For this word you have an experience of **Remembering** the word, this could have included seeing the word in your mind's eye, remembering what you thought or pictured when you saw the word on the original list, and/or having a sense of yourself in the past. *For example, if you see someone on the street you may think 'who is that? Oh yes, I remember, I was in the chemist shop, it's the person I saw in the queue at the chemist, I remember thinking what a funny hat they had on…'*

K = For this word you simply **Know** the word was on the previous list without any of the other feelings associated with vividly remembering that you have seen the word before. *For example, if you see someone on the street you may think 'who is that? Oh yes, it's my friend Rob, I know him really well…'*

F = For this word you have a feeling of **Familiarity** with the word and because of that you think that the word was on the previous list. *For example, if you see someone on the street you may think 'who is that? They look very familiar… I don't know where I know them from but they are definitely familiar…'*

G = For this word you had no feeling of familiarity or any other memories associated with the word and simply **Guessed** that the word had been on the previous list.

APPENDIX B: STIMULI FROM ALL EXPERIMENTS

B.1. CUE WORDS AND JUSTIFICATION STATEMENTS USED AS STIMULI IN EXPERIMENTS 2.1 TO 2.4.

The cue words and justification statements included in these experiments were those originally published by Gardiner et al. (1998). Experiment 2.1 involved participants making confidence judgments to all the justification statements. Items are sorted by their mean confidence value as obtained in Experiment 2.1 (from high to low within subjective experience category). Items were selected for inclusion in Experiments 2.2 and 2.3 on the basis of their confidence obtained in Experiment 2.1. Lowest, median, and highest values within subjective experience categories were chosen for Experiment 2.2, and eight items surrounding the mean confidence were selected for Experiment 2.3 (selections marked). Where items are not the exact lowest, median, highest, or mean item this is because the justification statement or cue word matched another included in that study and one had to be swapped for a neighbour, e.g., justification for *'Piano'* given a mean confidence of 92.05 (5th row of table) not included in Experiment 2.2 as another *'Piano'* already included (mean confidence 93.33, 2nd row of table). All Know and Familiar items were originally Know items according to Gardiner et al.'s participants and were split into Know and Familiar by expert raters HW and CM. All Know and Familiar items were used in Experiment 2.4.

Subjective Experience	Mean Confidence	Cue Word	Justification Statement	Item included in
Remember	94.77	Cranberry	I remembered it because when it came up it reminded me that we need more cranberry juice in the union bar where I work.	Exp. 2.2
Remember	93.33	Piano	When I saw it yesterday I remember saying. "Oh I can remember that because I can see myself playing the piano."	Exp. 2.2
Remember	93.13	Magazine	My friends laugh at me because I buy 5 magazines each week! This is the thought I had when I first saw the word.	

Subjective Experience	Mean Confidence	Cue Word	Justification Statement	Item included in
Remember	92.29	Hospital	I remembered it because of my father's condition; he is in a hospital. Seeing the word reminded me that I thought of that, it triggered a connection of thoughts.	Exp. 2.2
Remember	92.05	Piano	I remembered that there were a few musical instruments that came up and "piano" was one of those.	
Remember	91.61	Mother	I remember that word strongly, the word is still in my memory in the way I saw it yesterday.	Exp. 2.2
Remember	91.48	Kilt	I saw it yesterday for definite, I remember seeing the word on the screen.	
Remember	90.42	Athlete	I had an outline of a cartoon image of an athlete when I saw it yesterday, and it all popped back into my mind when I saw it today.	
Remember	90.36	Butterfly	I remember seeing it. Making an image and a picture of it.	
Remember	89.79	Sofa	A girl I know is called Sofia. When I saw the word yesterday it made me think of her.	
Remember	89.66	Tangerine	I remembered it. It was the second one up yesterday.	
Remember	89.46	Road	I have tried to relate the words to summer, I know I saw it because of the way I remember it, running down the road in the summer.	
Remember	89.00	Puppy	Because I am from Belgium, we call my father Puppy (nickname) so I could not forget that word.	
Remember	88.79	Mother	It was near the beginning; it was the second word.	
Remember	88.64	Boy	I remember it appearing on the screen.	
Remember	88.00	Gun	When I saw it yesterday I had an image of a gun and I thought it was a strange word to put in.	
Remember	87.95	Furniture	My friend was moving, I remember thinking of that.	

Subjective Experience	Mean Confidence	Cue Word	Justification Statement	Item included in
Remember	87.86	Piano	I used to play the piano. I thought of that when I saw the word yesterday.	
Remember	87.80	Tornado	I thought of an actual “tornado” when I saw it yesterday, so I remembered it today.	
Remember	87.60	Sauerkraut	I remembered it because I could not pronounce it yesterday!	
Remember	87.32	Emerald	I used a story: grasshopper was before it and I connected it with a green emerald.	
Remember	87.29	Pickle	I was thinking of cheese and pickle sandwiches when it came up yesterday.	
Remember	87.27	Horse	I remember thinking about my horse.	
Remember	87.27	Body	I could remember thinking that it was an easy word to remember.	
Remember	86.72	Sofa	I had an image of people sitting on it. I remembered it because I remember thinking about that image.	
Remember	85.80	Surf	It was one of the first words that came up, so I was trying to memorise all the first words, so I remember that one.	
Remember	85.69	Piano	I used to play the piano, so yesterday I thought of that, when I saw the word today I just remembered it from yesterday.	
Remember	85.37	Harp	I remembered when I saw the word I quickly saw a harp in a church. The same picture came back.	
Remember	85.34	Ape	I remembered that word with another word. I formed a picture of it. I knew as soon as I saw it because I remember thinking of the picture of monkeys at the time.	
Remember	85.21	Surf	I answered “Remember” because it reminded me of living in Plymouth when I saw the word yesterday, so I remembered it this morning.	
Remember	85.15	Blood	It conjured a strong image, so I remember seeing that word yesterday.	

Subjective Experience	Mean Confidence	Cue Word	Justification Statement	Item included in
Remember	85.00	Surf	It came up at the beginning and I associated it with “island”.	
Remember	84.64	Eye	It was one of the first ones yesterday.	
Remember	84.63	Kilt	It reminded me of going to Scotland when I saw it yesterday.	
Remember	84.33	Sauerkraut	I remember I had to look at it closely yesterday.	
Remember	83.40	Cranberry	Yesterday I was trying to use strategies to remember things. So I associated it with Delia Smith’s book, which has recipes full of cranberries.	
Remember	83.04	Car	I created all this story surrounding the word “policeman” and the car was part of it, I remembered the story.	
Remember	83.00	Grasshopper	Where I work the boss has a tarantula and locusts were fed to them, so I made that connection with grasshoppers.	
Remember	82.96	President	When I saw it on the screen yesterday I was thinking of presidents’ faces.	
Remember	82.88	Rectangle	Yesterday as it came on the screen I imagined a rectangle against an orange background and that image came back to me.	
Remember	82.86	Cranberry	It was the first word on the list.	
Remember	82.22	President	Yesterday I associated this word with the word “minister”. Today I automatically remembered about that association.	
Remember	82.12	Policeman	My brother is a policeman. I thought of that.	
Remember	82.12	Piano	The word “house” came before it, so I had an image of a house with a piano inside.	Exp. 2.2
Remember	82.08	Tangerine	I thought of food when I saw it and today that came back to me.	Exp. 2.2
Remember	82.05	Raspberry	I remember associating it with “apricot” and “plum”.	Exp. 2.2

Subjective Experience	Mean Confidence	Cue Word	Justification Statement	Item included in
Remember	82.04	Cigarette	Because I smoke. I remember seeing it yesterday because of the image that it conjured. The same thoughts came back to me yesterday.	Exp. 2.2
Remember	82.00	Green	I felt that when I saw it yesterday, I remembered that it was my favourite colour.	
Remember	81.85	Piano	I remembered it because “musician” came up first and also I have a friend who plays the piano so I linked the two together.	Exp. 2.3
Remember	81.67	Emerald	I recently bought a ring. I had this thought yesterday and it came back to me when I saw the word again.	Exp. 2.3
Remember	81.61	Broom	When this word appeared I remember saying to myself that I had to try and remember that word.	Exp. 2.3
Remember	81.55	Sauerkraut	It was probably the way it was spelled, I remember looking at it yesterday and saying; “What does that say!?”	Exp. 2.3
Remember	81.50	Policeman	I remembered that it closely followed the word “gun”, one came after the other.	Exp. 2.3
Remember	81.30	Butterfly	I thought of a Greek song about butterflies when I saw the word.	Exp. 2.3
Remember	80.56	Furniture	When I first saw it I thought about “chairs” and “tables”. That was brought back when I saw the word today.	Exp. 2.3
Remember	80.52	Bluebell	Yesterday it came up close to the word ring. I remembered that.	Exp. 2.3
Remember	80.36	Kilt	I remembered that I thought of a Scottish man.	
Remember	79.70	Cider	I associated it with “outsider” when I saw it yesterday, today I remembered the association.	
Remember	79.64	Piano	When that word came up I thought of me playing the piano.	
Remember	79.39	Kilt	I had a picture of someone standing on a hill in a kilt. That picture came back today.	

Subjective Experience	Mean Confidence	Cue Word	Justification Statement	Item included in
Remember	79.32	Magazine	This was also part of the story, I remembered an image of a magazine.	
Remember	79.17	Tornado	When I saw it yesterday it reminded me of the 'Rover' advert.	
Remember	79.00	Nun	I made a connection in my head: it came just after kids and the kids were on the sofa. Nun came after that.	
Remember	78.83	Harp	On Friday I was in a restaurant with a harpist. I remember thinking of that yesterday.	
Remember	78.79	Kite	When I read it yesterday I thought of another word, so that is what I remembered today.	
Remember	78.21	Orchid	It was close to the word "cider", so I made a connection between the two words.	
Remember	77.80	Rainbow	Same reason, I remembered a multicoloured telephone box outside the hospital.	
Remember	77.59	Surf	At the time I was counting the number of "S" words and I remembered that "surf" was one of them and it was short.	
Remember	77.59	Sofa	I remember picturing a sofa yesterday and relating it to furniture.	
Remember	77.59	Athlete	I had a picture of the Olympics.	
Remember	77.50	Sauerkraut	It reminded me of my German lessons. It looked odd as one of the words in the list.	
Remember	77.33	Ring	I was playing with my rings when I saw it yesterday.	
Remember	75.89	Gondola	Because I remember relating it to skiing, that thought came back to me as I saw it today.	
Remember	75.83	Leopard	Yesterday I was making associations when I saw the words, and I remembered those.	
Remember	75.45	Horse	I had an image of an horse.	
Remember	75.45	Harp	I remembered being present at a concert where music was playing.	

Subjective Experience	Mean Confidence	Cue Word	Justification Statement	Item included in
Remember	75.00	Grasshopper	Because I help this group of children, I thought about that yesterday and that stuck in my mind.	
Remember	74.81	Tornado	I remember watching a program on television about it. I thought of that yesterday, and today the word reminded me of the program.	
Remember	74.20	Hotel	My stepfather owns a hotel, yesterday I have tried to associate the words with people I knew and that is what came to my mind today.	
Remember	73.67	Cider	I saw a program on underage drinking recently; so I thought of that when I saw the word.	
Remember	73.52	Furniture	I thought of a chair when I saw it.	
Remember	72.83	Log	I was thinking of "logging" in to the computer.	
Remember	70.71	Green	I remember thinking of nature straightaway yesterday.	
Remember	70.19	Gondola	I remember saying to myself. "What on earth is that?!" when I saw it yesterday.	
Remember	70.18	Tornado	Yesterday I studied something about "hazards" so I associated the word with that.	
Remember	69.63	Blood	That word came almost directly after I created an image of the "musician" playing the "piano."	
Remember	66.61	Athlete	I remembered having a conversation with a friend of mine.	Exp. 2.2
Remember	60.52	Furniture	I think that when I was trying to remember the words, it was one of the first words that came up.	Exp. 2.2
Remember	59.63	Grasshopper	I think the word boulder was there and I imagined a boulder with a grasshopper.	Exp. 2.2
Remember	46.36	Ape	When it came up I remember thinking that there were lots of words with three letters.	Exp. 2.2
Know	88.57	Sauerkraut	I remembered it because it was an unusual word. I knew it was there.	Exp. 2.2

Subjective Experience	Mean Confidence	Cue Word	Justification Statement	Item included in
Know	85.45	Gun	I just knew that I knew it.	Exp. 2.2
Know	83.33	Bluebell	I am sure about that one, there were a couple of words which were similar and were part of the category flower.	Exp. 2.2
Know	79.55	Bluebell	I am sure about that one, there were a couple of words which were similar and where part of the category flower.	
Know	79.09	Sofa	I was just sure of seeing the word here and I knew I had seen it.	Exp. 2.2
Know	78.93	Gun	I remember it was there, I was sure it was there, it was familiar.	
Know	77.80	Magazine	One of a few words that I knew I had seen it before and it was familiar.	
Know	76.90	River	I felt I knew it was there as soon as I saw it.	
Know	72.41	Grasshopper	I just knew it was there, there were no thoughts associated.	
Know	69.39	Car	I just knew it was there, nothing came back to my mind.	
Know	67.78	Car	I just knew I had seen it yesterday, there was no story connected to it, it was just very familiar.	
Know	65.61	Cigarette	I just thought it was there, but again no story, but I knew it was there.	
Know	65.20	Crucifix	I was not 100% sure of a connection, but I was pretty sure it was there. It was definitely in my head.	
Know	63.57	Cigarette	I felt as if I have seen that word, I was sure that the word was there yesterday.	
Know	63.28	Athlete	I am sure I saw it, but I can’t remember why I think I saw it.	
Know	63.20	Piano	I am sure that it came up, but I do not remember anything in particular about it.	
Know	62.92	Grasshopper	I did not have any images, it did not come up with any images, I remember it just being there.	
Know	61.67	Gun	There was no association, I just had a feeling that I saw it, I was sure.	

Subjective Experience	Mean Confidence	Cue Word	Justification Statement	Item included in
Know	59.07	Tangerine	I recognised it as a word from yesterday, but I cannot really remember what I thought, I could not remember seeing it on the screen but I was sure it was there yesterday.	Exp. 2.3
Know	58.41	Orchid	I really just recognised without remembering it appearing yesterday.	Exp. 2.3
Know	58.13	Hotel	It is one of the words I thought it was there yesterday, but there was no particular reason . . .	Exp. 2.2
Know	57.86	Magazine	I am sure it was there but I could not place it.	Exp. 2.3
Know	57.05	Pickle	Same as paper, I do not remember seeing it, but I remember it was there.	Exp. 2.2 and Exp. 2.3
Know	56.25	Zipper	I just remember it coming up, it was an unusual word, it seemed it came up before, but I did not recall anything about it, no thought was associated with it.	Exp. 2.2 and Exp. 2.3
Know	54.17	Road	I felt I saw this word but could not find the exact feelings of when I saw the word, but I am sure I saw the word, I could though not find the background.	Exp. 2.2 and Exp. 2.3
Know	54.11	Log	I just thought I recognised it, I felt that it was there, but there was no story connected to it.	Exp. 2.3
Know	53.86	Shrimp	I knew it was there, I do not remember seeing it but it was familiar.	
Know	53.18	Church	I just thought it came up yesterday.	
Know	52.50	Cider	I was sure I had seen it, but I could not remember seeing it.	
Know	52.05	Rectangle	I am sure it was there but I do not remember thinking about it.	
Know	51.88	Cigarette	Pretty sure it was there, but I cannot go back to an image or anything that happened yesterday.	
Know	51.67	Nun	I could not remember the situation but I knew I saw the word somewhere.	
Know	51.40	Gondola	I just thought it was there.	

Subjective Experience	Mean Confidence	Cue Word	Justification Statement	Item included in
Know	51.30	Cranberry	It was an unusual word. I thought I recognised it as soon as I saw it. I could not remember thinking about it, but I thought I had seen it, I could not think where else I would have seen it.	
Know	50.36	Summer	Nothing connected to it, it was just part of the whole thing yesterday.	
Know	49.38	Sea	I thought I remembered it but there were no thoughts regarding the word, perhaps just the feeling of seeing it yesterday.	
Know	49.11	Furniture	I just felt sure that it was there but I could not remember seeing it at the time.	
Know	48.15	Road	In a similar way as for “river”, I was pretty certain, but I cannot fix in my mind when I saw the word.	
Know	48.13	Paper	It was as if I had seen it, I am sure it was there yesterday, but I cannot remember seeing it.	
Know	47.73	Cigarette	I just felt sure that it was there but I could not remember seeing it at the time.	
Know	47.41	Tangerine	I literally had a feeling that it was there yesterday but I could not remember anything else.	
Know	46.48	Church	I felt there was a strong familiarity with it, I thought I saw the word but it was not connected to anything.	
Know	41.79	Sea	It looked sort of as if I saw it, but there was nothing associated to it.	
Know	41.67	Sea	When I saw the word it was familiar, I just thought it was there but there was nothing else.	Exp. 2.2
Know	33.83	Body	A lot of words flashed very quickly yesterday and I was not always concentrating, so I thought I saw that one.	Exp. 2.2
Know	33.45	Grasshopper	I just thought that when I saw it I recognised it and I thought it was there, but it triggered nothing.	Exp. 2.2

Subjective Experience	Mean Confidence	Cue Word	Justification Statement	Item included in
Know	30.56	Ring	I think I remember seeing it, but there was no link or image. I can't remember feeling anything.	Exp. 2.2
Familiar	67.58	Summer	I knew I had seen it, I had thought about summer but I could not remember what my thoughts were, it was just a feeling of seeing the word.	Exp. 2.2
Familiar	62.08	Emerald	It looked familiar I remember maybe repeating it when I was memorising it, but I was not 100% sure.	Exp. 2.2
Familiar	61.00	Hospital	I have also been to a hospital recently so that could be why the word is so familiar to me, but I believed it was there.	
Familiar	60.68	Emerald	I remembered that there was a jewel, but I cannot remember when I saw the word.	Exp. 2.2
Familiar	58.39	Harp	I just felt that I saw the word yesterday. There was something about musical instruments.	
Familiar	57.86	Father	There was something about father, mother and child that seemed familiar.	
Familiar	57.50	Tablespoon	It was a familiar word to me.	Exp. 2.2
Familiar	56.20	Butterfly	It was one of those words that rang a bell.	
Familiar	56.06	River	I could imagined it written across the screen, I thought I might have seen it yesterday. I felt it was there, but I do not know why.	
Familiar	54.79	Professor	It was not in any of the little stories I made up to remember the words, but I had a strong feeling of familiarity.	
Familiar	54.33	Kite	Because it was something to do with the sky. I was fairly confident it was there but there was no direct connection.	
Familiar	54.09	Furniture	I just thought I had seen it, I recognised it and I was sure it was there. But also because I saw the word chair I thought maybe I was wrong but it was familiar.	
Familiar	53.28	Church	I felt as if I saw it yesterday.	

Subjective Experience	Mean Confidence	Cue Word	Justification Statement	Item included in
Familiar	51.17	Emerald	I think I remember linking it to something, but I cannot remember what to.	
Familiar	50.40	Boy	I reckon I saw it yesterday, I felt that I saw it yesterday, but I am not extremely sure.	
Familiar	50.19	Squirrel	I remembered something about squirrels, but I cannot remember what.	
Familiar	48.86	Summer	I could not remember it, but it sounded familiar, I think I did see it yesterday.	Exp. 2.3
Familiar	46.67	Rectangle	I think I saw this word (but maybe it was triangle!), I do not remember visualising a rectangle but I saw it.	Exp. 2.3
Familiar	46.60	Grasshopper	Like “cigarette”, no association only a feeling of familiarity.	Exp. 2.3
Familiar	46.07	Magazine	I was pretty sure it came out, but I could not identify the word.	Exp. 2.2 and Exp. 2.3
Familiar	45.54	Bluebell	I could remember that there was a flower, but I could not remember the specific flower.	Exp. 2.2 and Exp. 2.3
Familiar	44.67	Magazine	I had a feeling of familiarity.	Exp. 2.2 and Exp. 2.3
Familiar	44.63	Magazine	I think I remember seeing it, but it was not associated with anything, actually I was not sure whether it was “magazine” or “newspaper”.	Exp. 2.3
Familiar	43.33	Harbour	It was familiar, but I was confused. I knew it was there but could not be sure.	Exp. 2.2 and Exp. 2.3
Familiar	42.07	Island	It seemed familiar, I was not aware of other things.	
Familiar	41.38	Plum	I think I remembered it, but I was not sure as nothing came back to me about seeing it, but I knew it was there, it was sort of familiar.	

Subjective Experience	Mean Confidence	Cue Word	Justification Statement	Item included in
Familiar	40.86	Surf	I had a feeling that I had seen that word, I kind of remember the word: when it came up just before it reminded me that I had seen it somewhere.	
Familiar	40.00	Father	I thought I saw the word father with another word, but I was not sure.	
Familiar	40.00	Cranberry	It was just familiar.	
Familiar	39.83	Kilt	It seemed familiar but I wondered.	
Familiar	39.64	Sea	I remember something about “sea” but I was not sure whether it was related to “surf” or “sea”. I was not sure, but I remember thinking about sea, but did not know whether it was because of seeing the word sea or not.	
Familiar	38.93	Whale	I did not form any associations, it just seemed familiar.	
Familiar	38.89	Library	I think it came back to me but I cannot pinpoint actually seeing it.	
Familiar	37.96	Car	I had a feeling that it was there; I think car was there.	
Familiar	37.07	Boy	It is familiar because it is connected to my children. But I could not find any feelings or thought that I had yesterday that were associated with it.	
Familiar	36.07	Keg	It was a feeling that it was there. I was not sure whether it was at work or here that I came across that word.	
Familiar	36.07	Squirrel	I thought it was there, but no thoughts associated. Either that or we were talking about “squirrels” yesterday.	
Familiar	35.86	Record	I do think it had familiarity, I did not know whether it was because I had seen it before.	Exp. 2.2
Familiar	34.63	River	I answered “Know” because I was pretty sure it was there yesterday, but I couldn’t remember, I could not specifically remember seeing it.	
Familiar	34.63	Sea	I do not remember seeing it but it felt familiar.	

Subjective Experience	Mean Confidence	Cue Word	Justification Statement	Item included in
Familiar	32.78	Rectangle	It just sounded familiar, I couldn't remember anything about it I just had a vague idea that it was there yesterday.	Exp. 2.2
Familiar	30.83	Party	I think it came up but it triggered nothing.	Exp. 2.2
Familiar	30.54	Leopard	I do not remember seeing words of animals, but the word is familiar, I cannot place it exactly, I knew I read it but I could not relate it to anything.	Exp. 2.2
Guess	62.59	Flea	I am almost certain that it was there. But not entirely.	Exp. 2.2
Guess	57.50	Clarinet	It seemed the sort of word you would remember, so I just had a feeling it was one of those words.	Exp. 2.2
Guess	55.54	Gondola	It was the sound of the word that had some familiarity.	Exp. 2.2
Guess	55.00	Blossom	I felt the same as for "slipper", I saw this word recently.	Exp. 2.2
Guess	53.70	Island	I come from an island, so it could have been that I saw the word on the list or that the word was familiar because I come from an island.	
Guess	53.39	Cranberry	I thought that there were lots of fruit words yesterday, so I thought it was one of them.	
Guess	50.93	Clarinet	It seemed the sort of word you would remember, so I just had a feeling it was one of those words.	
Guess	50.30	Hotel	I remember that there were lots of things to do with holidays, or words to do with hotels, so I imagined that it was there.	
Guess	48.86	Road	I was not sure, but I thought it was in one of my connections.	
Guess	46.60	Holiday	I guessed that it was there as I made some connections with summer which I knew it was there, but no memories about it.	

Subjective Experience	Mean Confidence	Cue Word	Justification Statement	Item included in
Guess	46.48	Water	I remember that there were lots of things to do with sea, or words to do with water, so I imagined that it was there.	
Guess	45.40	Horse	I was not certain but I thought I recognised it.	
Guess	45.37	Uniform	When I saw the word I thought I recognised it, I had a feeling it might have been there.	
Guess	45.00	President	There was no real association, but I remember something about seeing the word “president” recently!	
Guess	44.33	Church	I think I tried picturing a church, but also someone was talking to me about churches the other day, so I might have confused the two things.	
Guess	44.20	Road	Same as above, it was possible that I have seen it before.	
Guess	44.14	Gondola	It seemed that I could recall an image of Venice, but I was not sure whether I was imagining it or whether the word was there.	
Guess	43.86	Holiday	I am eager to go on holiday so I am not sure whether I saw it here or whether I was thinking about it.	
Guess	43.62	Harp	I remember that yesterday I was thinking about “music”, so I took a guess that the word was there.	
Guess	43.18	President	There was no real association, but I remember something about seeing the word “president” recently!	
Guess	43.15	Eye	There were basically 3 groups of association, and this may not have been the key word, but maybe it was part of the association.	
Guess	42.32	Sky	It was a short word, so I believed that it was there as there were a number of short words.	
Guess	41.61	River	I could remember thinking of the word, it seemed as if I recollected it, but it was to vague to say I was sure.	

Subjective Experience	Mean Confidence	Cue Word	Justification Statement	Item included in
Guess	41.00	Furniture	It seemed familiar, but it was just me thinking of furniture in the room, or was it there? Also up to the point when "furniture" came up, I said I remembered all of them, so I thought they cannot all be right.	
Guess	40.80	Apricot	Again it looked kind of familiar.	
Guess	40.80	Cider	Again, likely to have been there.	
Guess	40.45	Paper	I thought that I might have seen it.	
Guess	40.00	Nun	I was not sure whether I had seen it or not, but I thought I did.	
Guess	39.83	Holiday	I am eager to go on holiday so I am not sure whether I saw it here or whether I was thinking about it.	
Guess	39.11	Body	I remember there was something to do with clothing, so I had a vague thought that it was there.	
Guess	38.93	Horse	I was not certain but I thought I recognised it.	
Guess	38.21	Tornado	I thought I recognised it, but I was not sure. I recognised it in the way that I thought it was more there than not. For example, with gun I knew that it was definitely not there.	
Guess	38.20	Raspberry	I was not sure whether it was me going to pick raspberries, or whether the word was there yesterday.	
Guess	36.59	Blood	I think it was there but I was not sure.	
Guess	36.30	Broom	I felt I had seen it, but I was not sure, I thought guessing was the appropriate answer.	
Guess	35.83	Thorn	It was not definitely in my head that it was there, but it seemed vaguely familiar.	
Guess	35.76	Water	I had a feeling that it was there but I was not sure.	Exp. 2.3
Guess	35.34	Squirrel	It just seemed as if it was one of the words that would have been there, I have quite a recent image of the word but I was not sure whether it was there yesterday.	Exp. 2.3

Subjective Experience	Mean Confidence	Cue Word	Justification Statement	Item included in
Guess	35.15	Gondola	I kind of remember it, but I did not know whether it was here that I saw it or elsewhere.	Exp. 2.3
Guess	34.79	Party	I think it was familiar but I was not sure.	Exp. 2.3
Guess	34.66	Claw	I thought that was there but I was not sure.	Exp. 2.3
Guess	34.40	Harp	It seemed that there were quite a few musical instruments, so I took a guess that it came up.	Exp. 2.2 and Exp. 2.3
Guess	34.32	Record	It could have possibly been there yesterday.	Exp. 2.3
Guess	34.29	Harbour	I lived by the sea all my life, so I was not sure whether I have encountered that word here or whether it is to do with home.	Exp. 2.3
Guess	34.24	Surf	I was not sure, I thought it was there.	Exp. 2.2
Guess	33.70	Rainbow	I think I saw some word similar to “rainbow” yesterday.	Exp. 2.2
Guess	33.13	Ring	I thought I recognised it but I was not sure. Because “ring” is familiar but I was not sure whether it was for that or whether I actually saw it.	Exp. 2.2
Guess	33.00	Furniture	There were a few long words, I thought it could have been one of them.	
Guess	32.78	Gun	It could have been there but I was not sure.	
Guess	32.71	Limousine	It seemed likely it could have been there, I could not remember it being there but I had a feeling that it might have been.	
Guess	32.42	Clown	I remembered an image but it could have been my imagination, I do not know whether it was there, so I just guessed.	
Guess	32.33	Father	I think I saw it but I was not sure.	
Guess	32.32	Camera	I thought it might have been there.	
Guess	32.05	Piano	It could have been there.	

Subjective Experience	Mean Confidence	Cue Word	Justification Statement	Item included in
Guess	31.67	Slipper	I saw this word somewhere recently at some point, but I am not sure whether it was there yesterday.	
Guess	31.21	Sea	I thought I saw it but I was not sure.	
Guess	31.11	Telephone	I was not sure whether I saw the word yesterday or I was thinking that I had to call someone.	
Guess	30.50	Clown	I remembered an image but it could have been my imagination, I do not know whether it was there, so I just guessed.	
Guess	30.42	Telephone	I was less sure that telephone was there than I was of butterfly for example.	
Guess	30.40	Zipper	It was just a guess. I guessed when I recognised things but I was not sure that they were in the list, with “know” I was sure it was there . . .	
Guess	30.36	Puppy	It was more the situation that I have heard that word pretty recently, but I was not sure whether it was here or not.	
Guess	30.00	Harp	It was a short word and an object, so maybe it could have been one of them.	
Guess	29.81	Ring	I seemed to keep pressing the “yes” button a lot. So I thought it was better saying that it was a guess, because of the 30%. But it seemed familiar.	
Guess	29.79	Library	I thought that was one of the words but I was not sure at all.	
Guess	29.55	Uniform	It was sort of familiar but I was not sure, maybe I had seen it somewhere else.	
Guess	28.57	Furniture	It seemed familiar, but it was just me thinking of furniture in the room, or was it there? Also up to the point when “furniture” came up, I said I remembered all of them, so I thought they cannot all be right.	
Guess	28.50	Car	I was not sure. I had a very vague feeling that perhaps I saw it.	

Subjective Experience	Mean Confidence	Cue Word	Justification Statement	Item included in
Guess	28.39	Library	The word was familiar to me but the familiarity was not very strong.	
Guess	27.22	Athlete	I was not sure, I kind of recognised it but I was not sure.	
Guess	27.04	Piano	I felt a that there wasn't such strong familiarity. Maybe I saw it, but there were no strong feelings.	
Guess	26.72	Hospital	I thought it might have been there but I was not sure.	
Guess	26.07	City	I was not sure.	
Guess	25.00	Harp	I was not sure about it, it could have been there just as well as it could have not been there.	
Guess	24.66	Library	I thought that was one of the words but I was not sure at all.	
Guess	24.14	Limousine	It could have been there, I was not certain whether it was or not.	
Guess	22.71	Furniture	It could have been one of the words, but I am not sure, so a guess seemed an appropriate response.	
Guess	22.32	Church	I kept saying "no" so I just guessed it was there because you said that 50% of the words were there.	
Guess	22.17	Leopard	Not sure.	
Guess	21.30	Puppy	I was not sure that the word was there.	
Guess	20.23	Car	I was not sure.	
Guess	20.00	Island	I was not sure if I saw it.	
Guess	20.00	Log	I was really guessing with that word.	
Guess	19.44	Bluebell	It was a pure guess, I had a very slight feeling that it was there. I am not sure whether I saw it yesterday or before.	
Guess	18.15	Rainbow	I really do not know, it was a guess.	
Guess	15.52	Letter	I was not sure whether it was there or not.	
Guess	14.29	Log	I was not sure, there was nothing associated.	
Guess	13.67	Officer	It was just a guess.	Exp. 2.2

Subjective Experience	Mean Confidence	Cue Word	Justification Statement	Item included in
Guess	12.24	Father	I kept saying “no” so I just guessed it was there because you said that 50% of the words were there.	Exp. 2.2
Guess	7.50	Shrimp	It was a total guess.	Exp. 2.2
Guess	7.22	Weed	I was not sure at all!	Exp. 2.2

B.2. FORENAME-SURNAME PAIRINGS USED IN EXPERIMENTS 3.1 AND 3.2.

Forenames were obtained from the Office of National Statistics (www.statistics.gov.uk) as the most popular given names in England and Wales in 2007. Surnames were taken from the National Health Service Central Register as the 40 most common surnames in England and Wales in 2007. Pairing was done pseudo-randomly, with matching initial letters (e.g., George Griffiths) avoided. Different spellings of the same name and unisex forenames (e.g., Alex) were excluded, and forename-like surnames (e.g., Thomas) were excluded as surnames. Matching of targets and lures was done so that the lure did not have the same initial letter as the target forename or the surname. Target and lure forenames were also matched so that there were approximately equal numbers (13 or 14) of target-lure pairs which did and did not match on gender.

Target Forename	Lure Forename	Surname
Grace	Molly	Clarke
Ruby	Dylan	Williams
Olivia	Max	Harris
Emily	Abigail	Patel
Jessica	Poppy	Brown
Chloe	Mohammed	Hall
Lilly	Harvey	Jackson
Ella	Millie	Davies
Amelia	Lewis	Green
Lucy	Amy	Mitchell
Charlotte	Isabelle	Johnson
Mia	Phoebe	White
Evie	Jasmine	Walker
Hannah	Katie	Jones
Megan	Erin	Roberts
Imogen	Scarlett	Baker
Madison	Leah	Allen
Elizabeth	Ava	Phillips
Sophia	Lucas	Watson
Keira	Isaac	Cook

Target Forename	Lure Forename	Surname
Jack	Adam	Edwards
Thomas	Jacob	Wood
Oliver	Ryan	Evans
Joshua	Liam	Hughes
Harry	Daisy	Robinson
Charlie	Jake	Moore
Daniel	Tyler	King
William	Holly	Thompson
James	Matthew	Cooper
Alfie	Emma	Wilson
Samuel	Caitlin	Wright
George	Alexander	Hill
Joseph	Freya	Smith
Benjamin	Luke	Turner
Ethan	Callum	Ward
Harrison	Connor	Parker
Jayden	Oscar	Price
Cameron	Edward	Bennett
Archie	Georgia	Young
Henry	Alice	Griffiths

B.3. Word lists used in Experiment 3.3.

Stimuli were 480 medium- to high-frequency words obtained from the MRC Psycholinguistic Database (mean familiarity rating of 545; range 488-652. Familiarity values refer to the printed frequency in the language and were derived from merging three sets of familiarity norms: Pavio, unpublished, Toglia and Battig, 1978; and Gilhooly and Logie, 1980). Words were limited to between four and seven letters in length and were pseudo-randomly assigned to list to achieve an equal alphabetical spread across lists. Counterbalancing of lists across tasks and versions is shown in Table B.3.1 (after the 12 lists).

List 1		List 2		List 3		List 4	
able	lecture	arrival	painter	alcohol	master	ability	lawn
active	near	boil	poor	answer	outcome	academy	limp
aisle	neutral	brake	rain	balloon	paper	alone	luxury
aware	panic	bullet	regular	blow	refined	assist	mirror
black	pattern	claim	rule	clock	romance	bark	misery
bean	plot	cloud	safe	close	sent	belt	moon
blossom	pork	courage	seat	control	similar	bitter	nurse
blush	praise	cowboy	second	dome	snail	broken	origin
brick	public	device	service	drink	sore	century	rapid
button	relax	error	side	fault	spruce	cost	shiny
cherry	rent	extent	sing	fertile	sugar	dancing	step
cigar	room	felt	slush	force	supply	eight	tall
content	runner	flame	speech	golden	tale	faint	tape
deal	saucer	gang	stain	here	theft	foreign	think
diving	send	gravel	support	humming	thought	fresh	travel
doorway	sheep	head	talent	idea	ticket	hardly	true
film	slight	history	thing	issue	tired	heavy	twist
guilt	stole	home	tight	lady	waist	incense	very
help	stumble	marry	value	lend	willing	knight	voice
kept	wind	naughty	whether	look	window	known	well

List 5		List 6		List 7		List 8	
after	mark	alter	jewel	being	past	bank	lunch
ammonia	message	banker	loyalty	broad	piano	biology	nation
argue	mind	basic	mild	calf	private	careful	plant
artist	moment	bleach	mouse	cause	receipt	crowd	pudding
band	nail	brain	mystery	choose	retain	crumb	rock
become	offend	brush	peddle	collar	rice	date	settled
been	owner	burn	peel	cover	roll	dream	sight
blue	pair	charity	pony	death	seek	family	singer
burner	pepper	chest	pound	dreamer	shell	flash	soar
capsule	permit	coil	repair	elbow	skill	frown	soul
coffin	potato	cotton	repeat	engine	smile	germ	spring
comedy	razor	crow	ring	figure	smooth	gray	stay
comfort	salary	diamond	rough	finish	song	grief	surface
dawn	short	duck	scab	frog	squeak	hook	them
extreme	tractor	else	search	gentle	stopped	horror	touch
filling	upon	fatigue	silence	hard	style	itch	trust
fire	vein	feet	slide	jersey	tempt	knee	truth
grammar	warm	fortune	spoke	lead	trial	left	twig
human	wise	gown	strict	library	unique	lighter	utensil
lane	worm	hurt	wood	mouth	weapon	loud	volume

List 9		List 10		List 11		List 12	
appeal	mature	about	love	anyone	land	above	made
aunt	merry	advice	modern	bang	mixture	account	match
bargain	nose	area	most	board	mood	also	menu
beat	note	avenue	nest	cafe	needle	bubble	moose
billed	passive	beach	notice	captain	once	clash	nickel
brat	singing	because	older	charm	patient	dare	park
builder	snap	boast	oven	common	piece	diet	pretty
cast	still	border	perhaps	cookie	place	disease	quickly
daisy	symbol	boulder	plate	cruiser	rare	eager	rate
destroy	test	bridge	pond	cure	reflex	easy	reptile
exhaust	tidy	carrot	prayer	dusty	safety	faster	rural
fight	toast	deck	rail	feel	shop	front	single
grape	ugly	deliver	relief	fish	slice	future	spider
inch	urban	drug	scent	form	some	garbage	steam
jelly	victory	farther	shrimp	further	succeed	half	task
just	wallet	from	steal	gravy	threw	hear	then
lemon	when	grown	wealth	hockey	tickle	heel	throw
lock	whisper	heap	week	ideal	toilet	limited	vivid
lovely	wreck	intense	wink	kick	wave	little	whisker
mankind	wrote	lion	yell	labour	weak	load	which

Table B.3.1. Counterbalancing of word lists across task (Yes/No or 2AFC) and word type (target/lure) to create eight versions. Key to blocks: A = Pre-Exposure with Yes/No test; B = No Pre-Exposure with Yes/No test; C = Pre-Exposure with 2AFC test; D = No Pre-Exposure with 2AFC test.

	Word List Number by Task and Target/Lure												
	Yes/No Task						2AFC Task						
	1st Block			2nd Block			3rd Block			4th Block			
Block Order	Pre-Exp	Target	Lure	Pre-Exp	Target	Lure	Pre-Exp	Target	Lure	Pre-Exp	Target	Lure	Version
A-B-C-D	1+3	1	3	9+10	2	4	5+7	5	7	11+12	6	8	1
	3+1	3	1	9+10	4	2	7+5	7	5	11+12	8	6	2
B-A-C-D	9+10	2	4	1+3	1	3	5+7	5	7	11+12	6	8	3
	9+10	4	2	3+1	3	1	7+5	7	5	11+12	8	6	4
A-B-D-C	1+3	1	3	9+10	2	4	11+12	6	8	5+7	5	7	5
	3+1	3	1	9+10	4	2	11+12	8	6	7+5	7	5	6
B-A-D-C	9+10	2	4	1+3	1	3	11+12	6	8	5+7	5	7	7
	9+10	4	2	3+1	3	1	11+12	8	6	7+5	7	5	8

B.4. Word lists used in Experiments 4.1, 4.2, and 4.3.

Stimuli were 128 medium-frequency words obtained from the MRC Psycholinguistic Database (mean familiarity rating of 424; range 350-480. Familiarity values refer to the printed frequency in the language and were derived from merging three sets of familiarity norms: Pavio, unpublished, Toglia and Battig, 1978; and Gilhooly and Logie, 1980). Words were limited to between five and eight letters in length and were pseudo-randomly assigned to list. Words always used as fillers are shown in *italics* at the end of each list.

List 1				List 2			
annex	frill	mechanic	shawl	alkali	dwelling	outset	thicket
archery	gallery	military	soprano	analogy	farewell	pancreas	torment
atrocity	giant	mistress	strut	angel	firewood	peasant	waltz
brownie	glare	monarch	syllable	armour	fleece	prelude	warrior
bureau	glimpse	prairie	tally	beast	fraud	prune	wiggle[37]
cleaver	havoc	prong	torture	booth	harness	remedy	willow
comrade	herdsman	rattle	upright	brook	imitator	reprisal	witch
crook	impetus	register	wiggle[37]	canal	imprint	retreat	wound
cynic	insight	rotation	*beloved*	castor	incline	ribbon	*freight*
dignity	janitor	saddle	*coral*	cavern	kennel	rogue	*infantry*
diner	lobby	scout	*garment*	chink	lodge	satchel	*maiden*
discord	lowland	scroll	*porch*	crane	mayor	shriek	*pendulum*
educator	lumber	sedative	*tunic*	defiance	merit	slumber	*sardine*
emulsion	meadow	shaker	*ballot*	dividend	mineral	spasm	*siren*
fielder	outbreak	tyrant	*creeper*	drove	olive	spike	*troop*
forelock	overlap	upheaval	*flock*	dullness	otter	sword	*slime*

[37] After data collection it was discovered that the word 'wiggle' had been on both word lists. All data for this item were therefore deleted.

APPENDIX C: FORMULAE FOR YES/NO AND 2AFC D-PRIME CALCULATIONS

The statistic *d'* from signal detection theory is used in recognition memory to measure how well participants can distinguish targets from lures, independent of bias. Different formulae are needed to calculate *d'* in Yes/No and 2AFC recognition as 2AFC is calculated to have a performance advantage of about √2 over Yes/No recognition. To compensate for this *d'* should be divided by √2 in 2AFC (Hacker & Ratcliff, 1979; Macmillan & Creelman, 2005). A correction was also introduced by Snodgrass and Corwin (1988) to allow *d'* to still be calculated if any of the component proportions were 1. In this correction, proportions of hits and FA (false alarms) are calculated by adding 0.5 to the raw scores, and 1 being added to the total number of items. This correction was used in all calculations of *d'* in this thesis.

The formula used to calculate *d'* in Experiments 4.1, 4.2, and 4.3 and the Yes/No blocks of Experiment 3.3 was:

$$d' = z\left(\frac{\text{hits}+0.5}{\text{items}+1}\right) - z\left(\frac{\text{FA}+0.5}{\text{items}+1}\right) \tag{1}$$

The *d'* formula used in the 2AFC recognition of Experiments 3.1, 3.2, and 3.3 was:

$$d' = \left(\frac{z\left(\frac{\text{hits } A+0.5}{A \text{ items}+1}\right) - z\left(\frac{\text{FA } B+0.5}{B \text{ items}+1}\right)}{\sqrt{2}}\right) \tag{2}$$

Where hits *A* was the number of *A* items correctly recognised and FA *B* was the number of false alarms made to *B* items; relationship shown in Table C.1.

Table C.1. Relationships between hits and FA in 2AFC *d'* calculations

	Response	
Correct Answer	Item A	Item B
Item A	Hits *A*	FA *A*
Item B	FA *B*	Hits *B*

Thanks to Kang Lee and John Wixted from the University of California, San Diego, for their website explaining SDT computations (Lee & Wixted, 2004).

APPENDIX D: FORMULAE FOR COMPARING CORRELATIONS

To compare correlation coefficients the Fisher (1921) *r*-to-*z* transformation is commonly used to transform *r* values, which may not have a normal distribution, into *z* values, which are normally distributed. The formula for Fisher's transform is:

$$z_{jk} = ½\, ln\left(\frac{1+r_{jk}}{1-r_{jk}}\right) \tag{3}$$

Where *j* and *k* are the two variables you are correlating and *ln* is the natural logarithm.

When comparing coefficients these *z* values can be used in different formulae depending on whether the coefficients you are comparing come from the same sample or different samples. To compare whether two coefficients from different samples are significantly different Fisher's *Z* can be calculated using the Fisher's transformed coefficients z_{jk} and z_{hm} and their corresponding *N* values. This formula was used in Experiment 3.1:

$$Z = \frac{(z_{jk} - z_{hm})}{\sqrt{\left(\frac{1}{N_{jk} - 3}\right) + \left(\frac{1}{N_{hm} - 3}\right)}} \tag{4}$$

Steiger (1980) compared a variety of formulae which could be used to test whether two correlations were significantly different when the correlations were not independent, i.e., they were obtained from the same sample. He concluded that the $\bar{Z}_2$* test employing Fisher-transformed correlations was the best test for the comparison of two correlations which do not have a factor in common. This formula was used in Experiments 3.2 and 3.3:

$$\bar{z}_2 * = (N - 3)^{½}(z_{jk} - z_{hm})\,(2 - 2\bar{s}_{jk,hm})^{-½} \tag{5}$$

Where $\bar{s}_{jk,hm}$ is the pooled sample estimate of r_{jk} and r_{hm}.

Thanks to Chris Fife-Shaw from the University of Surrey for putting these formulae in a very easy-to-follow Excel document on his website and making it available for all to use (Fife-Shaw, 2005).

APPENDIX E: FORMULA FOR GAMMA CORRELATION

The gamma correlation is a non-parametric test of association (Goodman & Kruskal, 1954). It is typically used in metacognition research where the association is between predicted memory performance and observed memory performance (Nelson, 1984). For each participant a matrix like Table E.1 is produced, comparing participants' predictions with the number of times they retrieve or fail to retrieve an item at each level of prediction.

Table E.1. The gamma correlation with two response categories

	Memory Performance	
Prediction of Performance	Retrieved	Not Retrieved
Will retrieve	a	b
Won't retrieve	c	d

Where the values of a, b, c, and d are frequencies.

To calculate gamma (γ):

$$\gamma = \frac{ad-bc}{ad+bc} \qquad (6)$$

In Experiment 2.4, instead of the association being between predicted performance and actual performance, it was between experimenter ratings of Know and Familiar and participants' decisions as to which items were Know or Familiar items. For each participant a table like Table E.2 was produced and a gamma value calculated using the above formula.

Table E.2. The gamma correlation used in Experiment 2.4.

	Participant's Response	
Expert Classification	Know	Familiar
Know	a	b
Familiar	c	d

APPENDIX F: DECISION CRITERIA QUESTIONNAIRE FOR EXPERIMENT 2.4

Decision Criteria Questionnaire – Card Sorting Task

Nb. On each page please complete all the questions before reading the questions on the next page.

What are your initials? ______________ (the same as you wrote on your cards)

In this study we were interested in whether you thought there were any differences in memory justification statements. You were given 90 cards on which were written 90 cue words and statements and asked to sort these into Type A and Type B statements.

We are now interested in what criteria you decided to sort the statements on. In what way(s) did you think that the statements that you decided were Type A differed from those you decided were Type B? In the box below please write in as much detail as you can, in your own words, how your Type A and Type B statements differed. Give examples from the cards if you want to.

How easy did you find it to sort the statements?

☐	☐	☐	☐	☐
Very easy	Easy	Neither difficult nor easy	Difficult	Very difficult

How similar or dissimilar did you think Type A and Type B statements were?

☐	☐	☐	☐	☐
Very similar	Quite similar	Neither similar or dissimilar	Quite dissimilar	Very dissimilar

How likely do you think it is that other people would sort the statements in the same way as you have?

☐	☐	☐	☐	☐
Very likely	Likely	Neither likely nor unlikely	Unlikely	Very unlikely

When you were sorting the statements into Type A and Type B statements which, if any, of the following reasons influenced your sorting decisions *(please tick all that influenced you):*

- ☐ I thought one Type of memory statement used more emotional language than the other Type
- ☐ I thought one Type of memory statement included more visual imagery than the other Type
- ☐ I thought that for one Type of memory statement the participants who had made those statements must have been more confident in their memory performance than the other Type
- ☐ I thought one type of memory statement used more abstract words than the other Type
- ☐ I was mainly guessing
- ☐ I thought one Type of memory statement had used a deeper level of processing than the other Type
- ☐ I thought one Type of memory statement had included more concrete words than the other Type
- ☐ I thought that for one Type of memory statement the people making the statements had recalled more information in their memories than for the other Type of memory statement
- ☐ I sorted the statements on gut instinct
- ☐ I thought that for one Type of memory statement the people making the statements sounded more sure of their memories than for the other Type of memory statement

Again, please ask if you do not understand any of the questions on this questionnaire, or you need something explaining.

(Please turn the page)

For the statements that you have labelled Type A, how confident in their memory performance do you think the participants were who made those statements?

☐ % confident (0=low, 100=high)

And for the statements that you have labelled Type B, how confident in their memory performance do you think the participants were who made those statements?

☐ % confident (0=low, 100=high)

For the statements that you have labelled Type A, how much information or how many details do you think the participants recalled in their memories for the cue words?

☐	☐	☐	☐	☐
A great amount of info/detail	Quite a lot of info/detail	A medium amount of info/detail	Not much info/detail	No info/details

And for the statements that you have labelled Type B, how much information or how many details do you think the participants recalled in their memories for the cue words?

☐	☐	☐	☐	☐
A great amount of info/detail	Quite a lot of info/detail	A medium amount of info/detail	Not much info/detail	No info/details

(Please turn the page)

Below are two definitions which are typically provided to participants in memory experiments. Participants would use these statements after they had responded YES - they did think a particular item had been on the list of items they had had to learn. We are interested in whether these definitions fit with your definitions of Type A and Type B memory statements?

K = For this item you simply **Know** that the item was on the previous list without any of the other feelings associated with vividly remembering that you have seen the item before. *For example, if you see someone on the street you may think 'who is that? Oh yes, it's my friend Rob, I know him really well…'*

F = For this item you have a feeling of **Familiarity** with the item and because of that you think that the item was on the previous list. *For example, if you see someone on the street you may think 'who is that? They look very familiar… I don't know where I know them from but they are definitely familiar…'*

The definition of K fits my Type _______ (A or B) statement….

☐	☐	☐	☐	☐
Extremely well	Well	Quite well	Only slightly	Not at all

The definition of F fits my Type _______ (A or B) statement….

☐	☐	☐	☐	☐
Extremely well	Well	Quite well	Only slightly	Not at all

If you have any other comments to make about this experiment please write them here:

Thank you for completing this experiment, please take this questionnaire and your cards back to the experimenter.

APPENDIX G: ITEMS ANALYSIS, EXPERIMENT 2.4

To further examine the different results obtained for Know and Familiar items, the proportion of participants who had assigned each item to the correct category was calculated. From Table G.1 we can see that the mean proportion of items correctly assigned matches the means calculated by participant, as shown in Section 2.5.3.3. By conducting the items analysis however we can also see that the proportion of items correctly categorised had a larger range for Familiar compared to Know justifications, and the minimum and maximum proportion of items correctly categorised were also lower for Familiar compared to Know justifications. This suggests that there was more consistency in categorisation of Know justifications compared to Familiar and fits with the finding that Know justifications were able to be categorised correctly whereas Familiar judgments were not.

Table G.1. Descriptive statistics for Know and Familiar items.

Justification Type	N	Min.	Max.	Range	Mean	SD
Know	47	.45	.94	.49	.74	.12
Familiar	43	.20	.80	.59	.52	.15

APPENDIX H: ALTERNATIVES TO FIGURES 3.8 AND 3.9, EXPERIMENT 3.3

When compared to their originals in Chapter 3, these alternatives: Figures H.1 and H.2, which include all available data, demonstrate that the reduction in N due to listwise exclusion of data in the ANOVAs did not lead to a change in the patterns of means.

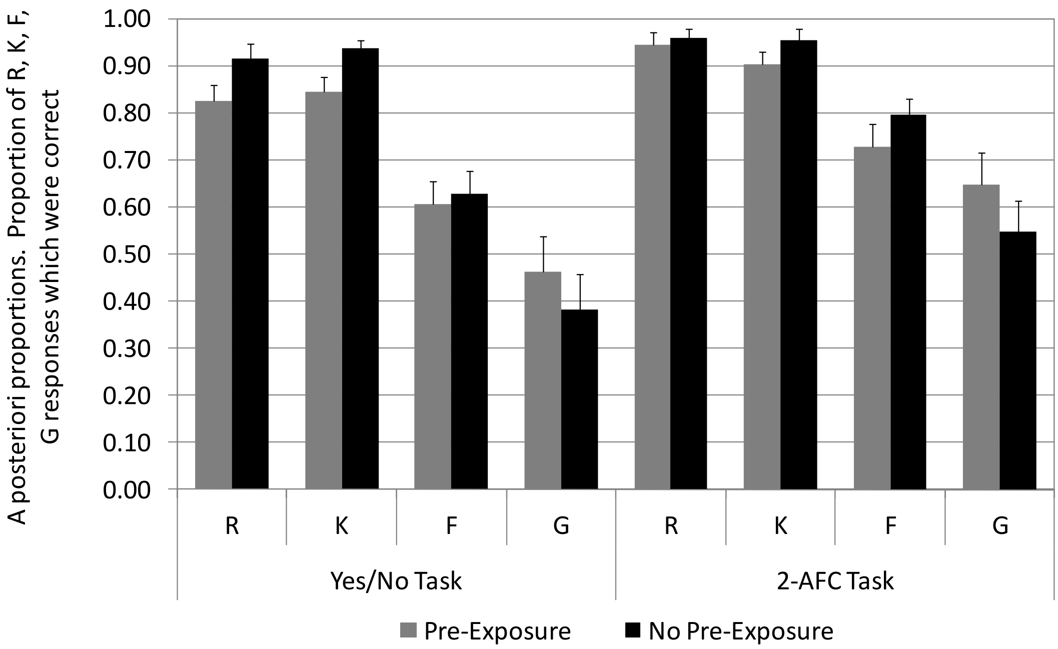

Figure H.1. Alternative to Figure 3.8. Proportion of items assigned to Remember (R), Know (K), Familiar (F), and Guess (G) in the in the Pre-Exposure and No Pre-Exposure conditions of the Yes/No task and the 2AFC task that had previously been correctly recognised. Errors bars = 1 SeM.

Table H.1. Number of participants included in calculating the means displayed in Figure H.1 above.

	Yes/No Task				2AFC Task			
Condition	R	K	F	G	R	K	F	G
Pre-Exposure	36	31	34	26	29	27	28	19
No Pre-Exposure	37	31	33	20	32	28	25	24

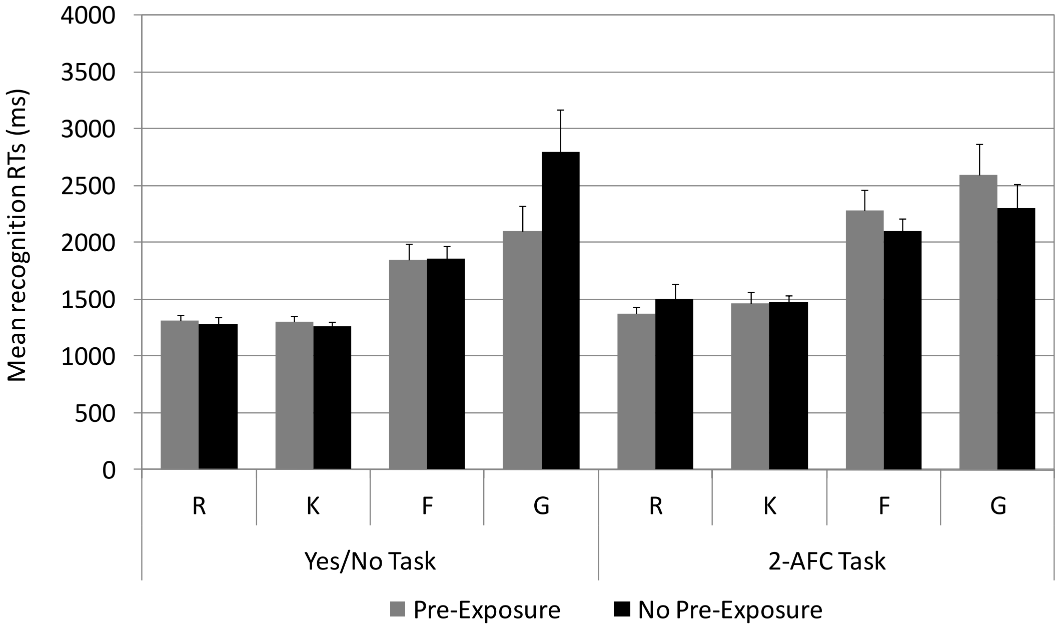

Figure H.2. Alternative to Figure 3.9. Mean correct recognition RTs by later subjective experience category, Remember (R), Know (K), Familiar (F), and Guess (G), in the in the Pre-Exposure and No Pre-Exposure conditions of the Yes/No task and the 2AFC task. Errors bars = 1 SeM.

Table H.2. Number of participants included in calculating the means displayed in Figure H.2.

	Yes/No Task				2AFC Task			
Condition	R	K	F	G	R	K	F	G
Pre-Exposure	35	31	32	18	29	27	27	17
No Pre-Exposure	36	31	31	13	32	28	25	20

APPENDIX I: ALTERNATIVE TO FIGURE 4.15, EXPERIMENT 4.2

When compared to the original Figure 4.14 in Chapter 4, this alternative, which includes all available data, demonstrates that the reduction in N due to listwise exclusion of data in the ANOVAs did not lead to a change in the patterns of means.

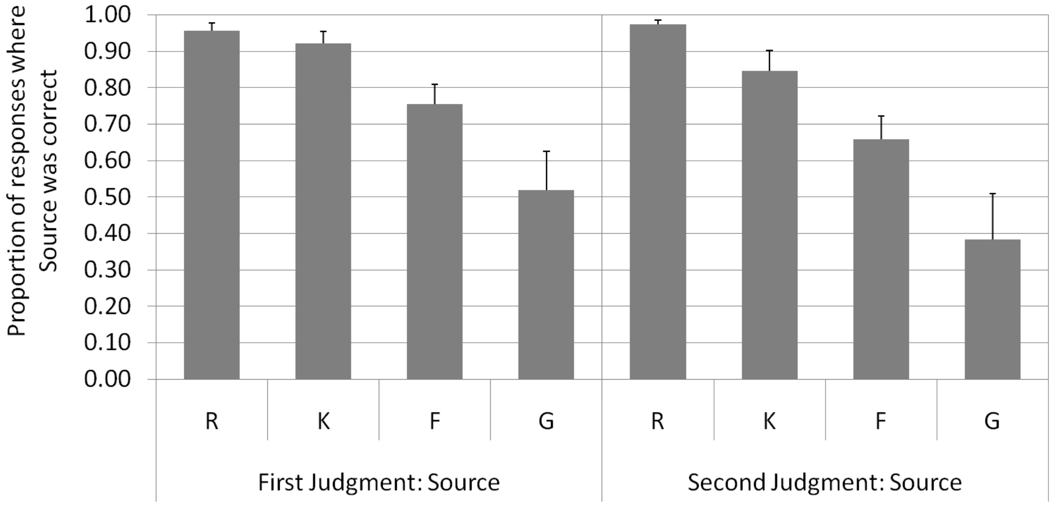

Figure I.1. Alternative to Figure 4.15, Experiment 4.2. Left panel: Of the items assigned to each subjective experience category at Second Judgment, the mean proportion given correct source judgments at First Judgment. Right panel: Of items assigned to each subjective experience category at First Judgment, mean proportion later given correct source judgments at Second Judgment. Error bars = 1 SeM.

Table I.1. Number of participants included in calculating the means displayed in Figure I.1 above.

Left panel (First Judgment: Source)				Right panel (Second Judgment: Source)			
R	K	F	G	R	K	F	G
29	25	31	14	30	27	30	12